Characterisation of hyaluronate lyases from streptococcal species

Anna-Marie Lindsay

A thesis submitted in partial fulfilment of the requirements of the University of Northumbria for the degree of Doctor of Philosophy

Research was undertaken in the School of Applied Sciences and in collaboration with York Structural Biology Laboratory, York University and The Animal Health Trust, Suffolk.

September 2007

Abstract

Previously cloned bacteriophage encoded HylP1 from the genome sequenced organism *Streptococcus pyogenes* SF370 was expressed in *Escherichia coli* and purified to homogeneity. The protein, previously assigned to glycoside hydrolase family 69 (GH69) was biochemically recharacterised as a polysaccharide lyase and reassigned to family 16 (PL16). The enzyme demonstrated a K_m of 1.47 mg ml^{-1} and a k_{cat} of 7.2 s^{-1}. Biochemically the enzyme had an optimum pH of 6.5 and temperature of 37 °C. The enzyme required no additional divalent ions for catalysis. The enzyme demonstrated strict substrate specificity only degrading hyaluronate and with no activity against related substrates chondroitin 4 sulphate and chondroitin 6 sulphate. HPAEC indicated the HylP1 had an endo mechanism of cleavage producing a range of differently sized oligosaccharides with the smallest being a tetramer. Site directed mutagenesis revealed a role for residues D157 and Y169 with substitution of these residues with alanine resulted in a 88.5% and 91.9 % loss of activity respectively. The location of these residues within the solved structure of HylP1 falls within the triple stranded β helix formed by the trimerised protein. This region of the protein was cloned, expressed and characterised and demonstrated similar kinetics as the full length protein (K_m of 0.53 mg ml^{-1} and k_{cat} of 11.1 s^{-1}). The activity of the enzyme when compared to other hyaluronate lyases shows it to be relatively inefficient yet when compared to other bacteriophage encoded hyaluronate lyases, HylP1 was very similar. The proposed role of these bacteriophage encoded hyaluronate lyases is one of degradation of the hyaluronate capsule surrounding the streptococcal cells to allow for penetration of the bacteriophage during infection.

Using the sequence of HylP1 the recently completed genome of *Streptococcus equi* was searched using bioinformatics tool BLAST. This revealed the presence of a protein, SEQ2045, sharing 85 % identity with HylP1. The protein was cloned and expressed in *E. coli* and biochemically characterised as a hyaluronate lyase. The enzyme demonstrated a K_m of 2.05 mg ml^{-1} and a k_{cat} of 6.2 s^{-1} which when compared to those of HylP1 is suggestive that the two enzymes are strongly related. The enzyme had an optimum pH of 6.5 and temperature of 37 °C and like HylP1 demonstrated only activity against hyaluronate and had an endo mechanism of cleavage with the smallest product of digestion being a tetramer. Site directed mutagenesis of the same residues as in HylP1 again yielded reduced activity (91.3 % and 87.6 % respectively).

Bioinformatic analysis of the genome of *S. equi* was performed by BLAST searching with the proposed gene sequences of *S. equi* flanking SEQ2045. This allowed for the production of a prophage map which shows distinct similarities to the prophage map of *S. pyogenes* suggesting both may be of the same origin.

Purified SEQ2045 was used in western blot analysis with *S. equi* convalescent horse serum. A strong positive reaction demonstrated a possible role for SEQ2045 during an infection suggesting that the bacteria have acquired this enzyme as a potential virulence factor by horizontal gene transfer. This presents a useful opportunity for the study of both *S. pyogenes* and *S. equi* infection process and the role of bacteriophage by the use of *S. equi* as a model for *S. pyogenes*.

Contents

Table of Figures

Table of tables

Acknowledgements

I would like to thank firstly the University of Northumbria for the opportunity to complete my PhD and Professor Gary Black, Professor Iain Sutcliffe and Dr. Simon Charnock for their advice and support. I would also like to thank Professor Andrew Waller from the Animal Health Trust and Dr. Edward Taylor from the York Structural Biology Laboratory at York University.

I would also like to thank all members of lab A307/9 past and present for their support and encouragement throughout.

Finally, I would like to thank my family and Justin for the emotional support they've provided and the belief in me which they have shown throughout. It is this, that has pushed me onwards to achieve anything that is possible.

"Anybody who has been seriously engaged in scientific work of any kind realizes that over the entrance to the gates of the temple of science are written the words: 'Ye must have faith.' It is a quality which the scientist cannot dispense with."

Max Planck

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. The work was done in collaboration with York Structural Biology Laboratory, York University and The Animal Health Trust, Suffolk.

Name: Anna-Marie Lindsay

Signature: 

Date: 05/06/08

Abbreviations

A	Adenosine (DNA)/Alanine (protein)
A_x	Absorbance at x nm
APS	Ammonium persulphate
ATP	Adenosine triphosphate
bp	Base pair(s)
BSA	Bovine serum albumin
C	Cytosine
°C	Degree Celsius
CBM	Carbohydrate binding module
CFE	Cell free extract
C-terminal	Carboxyl terminal
Da	Dalton
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
DNSA	Dinitrosalicylic acid
dNTP	deoxynucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylene diamine tetraacetic acid
F	Phenylalanine
g	Gram(s)
G	Gaunidine
GAG	Glycosaminoglycan
GH	Glycoside hydrolase
GT	Glycosyltransferase
h	Hour(s)
HEPES	N-[2-hydroxyethyl]piperazine-N'-[2-ethanosulphonic acid]
His_6 tag	Hexahistidine tag
HPAEC	High performance anion exchange chromatography
I	Isoleucine
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-β-D-Thiogalactopyranoside
kb	Kilobase pair(s)
kcat	Turnover number
kDa	Kilodalton
Km	Michaelis-Menten constant
l	Litre(s)
LB	Luria-Bertani medium
M	Molar
mA	Milliamps
min	Minute(s)
mm	Millimetre(s)
MW	Molecular weight
N-terminal	Amino terminal
OD_x	Optical density at x mn
ORF	Open reading frame
p	Plasmid

pDNA	Plasmid DNA
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PL	Polysaccharide lyase
Psi	Pounds per square inch
rpm	revolutions per minute
s	second(s)
SDM	Site directed mutagenesis
SDS	Sodium dodecyl sulphate
sp.	Species
subsp.	Subspecies
ssDNA	Single stranded DNA
TAE	Tris-Acetate-EDTA
TEMED	N,N,N',N'-tetramethylethylene diamine
T7 *lac*	lac operator just downstream of T7 promoter
T_m	Melting temperature
Tris	tris(hydroxymethyl)aminomethane
UV	Ultraviolet
V	Volts
V_{max}	maximum velocity
v/v	volume per volume
w/v	weight per volume
x g	times gravity
Y	Tyrosine
ε	Molar absorptivity
α	Alpha
β	Beta
K	Kilo
λ	Lambda
m	milli
μ	micro
1^o	Primary
2^o	Secondary
3'	Three prime
5'	Five prime
18.2 MΩ cm^{-1} H_20	18.2 mega ohm water

1 Introduction

1.1 Connective tissue

Connective tissue is essential for the maintenance of the healthy structure and function of the animal body. It can be classified into two distinct groups; connective tissue proper (which includes loose connective tissue, adipose tissue and dense connective tissue) and specialised connective tissue (which includes bone, blood and cartilage)

1.1.1 Specialised connective tissue

1.1.1.1 Cartilage

Normal articular cartilage is the connective tissue found covering the ends of long bones in synovial joints and it is unique within connective tissue as it has no blood vessels, lymphatic vessels or nerve fibres and lacks the separating basement membrane on either side of the tissue seen in other cartilage (Kuettner, 1992). Cartilage is comprised of two main components: the tissue fluid and the complex framework of three structural macromolecules that gives a tissue its form and stability. The structural macromolecules found in cartilage: collagen, proteoglycans and non-collagenous proteins, contribute between 20-40 % of the wet weight of the tissue (Buckwalter and Mankin, 1998) with the remainder of the wet weight of articular cartilage being water which is tightly bound within the extracellular matrix (Figure 1.1).

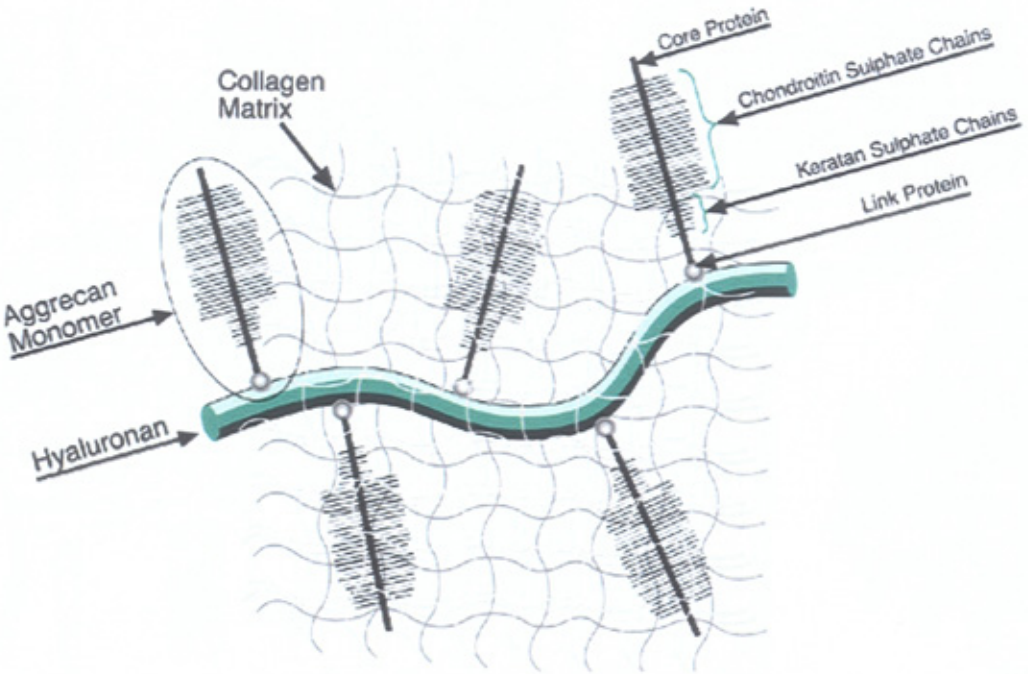


Figure 1.1: The structure of articular cartilage
http://www.peprotech.co.kr/fa_sub/img/chart-2.jpg

Articular cartilage matrix has biomechanical properties provided by the unique structural arrangement of the macromolecules, such as tensile strength and elasticity which facilitate both the absorption and distribution of loads placed upon the joint (Huber et al., 2000). Within the macromolecular arrangement the collagen fibres are arranged into an ordered three-dimensional (3D) network providing tensile strength and maintainence of tissue volume and shape. The proteoglycans which are highly hydrated due to the overall negative charge of the molecule attribute the cartilage with the ability to be reversibly deformed and resist any external compressive forces. It is these properties that make proteoglycans crucial in the distribution of load in weight-bearing joints (Schiller et al., 2006). Proteoglycans are comprised of protein core with one or more covalently attached glycosaminoglycans (see section 1.2). The glycosaminoglycans are long polysaccharides comprised of repeating disaccharide subunits and they confer to the proteoglycan a net negative charge which is responsible for the biomechanical properties observed of it. The net negative charge of proteoglycans will attract water and provide a swelling pressure within the matrix. Under regular conditions, the tight and ordered

collagen network limits this swelling of the proteoglycans (Kuettner and Thonar, 1998). Under load bearing conditions, the complete cartilage matrix undergoes deformation due to the fluid being expelled away from the loaded region towards the synovial cavity. This movement moves water away from the proteoglycans within the matrix. As a result, the proteoglycans become forced closer together which has the effect of increasing the negative charge of the matrix. This increase in negative charge creates an intermolecular charge repulsion, which serves to greatly increase the resistance of the tissue to further deformation. The deformation finally reaches an equilibrium point where the external loading force becomes balanced by the internal swelling pressure forces. When the load is removed the expelled water returns from the synovial cavity, carrying with it nutrients for the cells.

1.2 Glycosaminoglycans

Glycosaminoglycans (GAGs) are long polysaccharides consisting of repeating disaccharide subunits. This subunit consists of an N-acetyl hexosamine and a hexose or hexuronic acid. The members of the GAG family differ in the variety of subunit constituents and the glycosidic link between the subunits. The GAG family is composed of hyaluronan, chondroitin, heparan, dermatan and keratan. GAGs except hyaluronan can be covalently linked to a protein core to form what is known as a proteoglycan. Both structures are found in connective tissue and the extracellular matrix. They are found widely in the body including the extracellular matrix, the vitreous humor of the eye and the Whartons jelly of the umbilical cord. In rheumatoid arthritis the levels and biological nature of the GAGs changes, with an overall increase in their concentration. An increase in the amount of GAGs in the synovial fluid is also a marker of joint destruction (Wang and Roehrl, 2002). Changes in composition and distribution of GAGs is also proposed to be associated with the inflammation of the intestinal mucosa in Crohn's disease (Belmiro et al., 2005). As well as being associated with inflammatory and destructive diseases, GAGs have also been observed to play a role in the limitation of oxidative injury during progression of diseases yielding free radicals. An increase of GAGs during iron induced oxidative injury has been

shown to reduce cell death suggesting that the molecules function as antioxidants (Campo et al., 2005). GAG binding sites have also been observed on the viral envelope on a wide range of viruses including flaviviruses (hepatitis C, yellow fever) and herpes simplex (Olofsson and Bergstrom, 2005, Olenina et al., 2005).

1.2.1 Hyaluronan

Hyaluronan was first biochemically purified by Karl Meyer in 1934 from the vitreous humor of the eye and its structure of repeating disaccharide subunits of glucuronic acid and *N*-acetyl glucosamine (Figure 1.2) was then solved in 1954 (Lee and Spicer, 2000). It is unique from other GAGs in that it is not sulphated at any point in humans. It is synthesised in the plasma membrane by 1 of 3 different forms of the enzyme hyaluronan synthase (HAS1, HAS2 and HAS 3) (Adams et al., 2006) with elongation of the molecule occurring by the addition of alternating UDP glucuronic acid and UDP *N*-acetyl glucosamine molecules at the reducing end of the growing polysaccharide and as the chain increases in length the polymer is extruded through to the outside of the cell (Mummert, 2005). Unlike other GAG chains, hyaluronan is not bound to a protein core in the form of proteoglycan but remains free although some isolations of hyaluronan do yield small amounts of protein (Vynios et al., 2002). Hyaluronan is the largest of the GAGs with each chain being comprised of approximately 2500 disaccharide units (Vynios et al., 2002) giving an average molecular weight of 2×10^4 kDa (Stern et al., 2006). The structure of the hyaluronan chain shows that each disaccharide unit is twisted about 180° compared to the previous and next members of the chain. This twisting results in the entire chain being twisted 360° every two disaccharide subunits causing a twofold helix (Bayraktar et al., 2004). The carboxylate group of the hexauronic acid residue in each disaccharide unit confers a net negative charge on the molecule which then allows it to act as a soluble cation exchanger molecule contributing to the regulation of ion movement through the extracellular matrix in which it is found (Bayraktar et al., 2004). A number of hyaluronan binding proteins have been discovered indicating that hyaluronan is not just a packing material but serves an important role in

physiological processes. These binding proteins all have a specific binding sequence and as such are classed into a group called the hyaladherins. They can be either matrix associated or cell bound (Fraser et al., 1997).

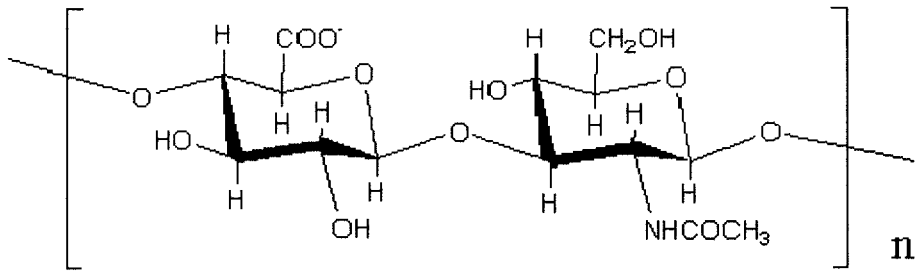


Figure 1.2: Structure of hyaluronan. Hyaluronan is a glycosaminoglycan (GAG) composed of a repeating disaccharide subunits with the structure (β-1,3-N-acetyl-D-glucosamine-β-1,4- D-glucuronic acid)n. (Taken from http://www.sci-museum.kita.osaka.jp/~takegawa/k-hyaluronan/hyaluronan.gif)

1.2.1.1 Hyaladherins

1.2.1.1.1 Matrix associated hyaladherins

Aggrecan is a chondroitin sulphate proteoglycan found in the extracellular matrix and is the most common proteoglycan of articular cartilage (Seyfried et al., 2005). It binds to hyaluronan via a small glycoprotein, link protein, to form a large highly hydrated aggregate molecule which contributes 35 % of cartilage dry weight and provides cartilage with compressive resistance (Dudhia, 2005). The high level of hydration provides the aggregate with increased load bearing and lubrication properties for articular cartilage. Destruction of either the aggrecan or the hyaluronan aggregates causes a decrease in hydration and as such can lead to degenerative disorders of the joints such as osteoarthritis (Chockalingam et al., 2004, Rodriguez and Roughley, 2006).

1.2.1.1.2 Cell bound hyaladherins

Cell bound hyaladherins are receptors for hyaluronan and the response of different receptors differ depending on the location of the cell and the cell type. The receptor for hyaluronic acid mediated motility (RHAMM) is a receptor which is not just surface bound but is also found in the cytoplasm of cells and can be secreted in a soluble form by cells. Upon binding hyaluronan its principal role is migratory but evidence suggests it is also involved in signal transduction within the cell following binding and in control of the cell cycle (Fieber et al., 1999). Hyaluronan can also bind to the widespread CD44 molecule found on cell surfaces. CD44 molecule binding to hyaluronan has been shown to cause cell migration, activation of leukocytes and lymphocyte homing (Schultz et al., 2005).

1.2.1.2 Functions of Hyaluronan

Following its discovery, hyaluronan was generally regarded as simply a space packing material with no physiological function. Following studies it has since become apparent that not only does it form the extracellular matrix, it also plays an important role in many physiological processes including cell proliferation and migration, inflammation, cell differentiation and tumour development and also has an immunological role and provides a protective mechanism against oxidative stress as well as its role in the formation and maintenance of the extracellular matrix.

1.2.1.2.1 Cell proliferation

Various studies have reported on the role of hyaluronan in cell proliferation and migration. Determination of the gene expression of hyaluronan related enzymes has shown that in proliferating cells there is a marked increase in the hyaluronan synthase genes and in non-proliferating cells an increase of hyaluronan degrading enzymes, hyaluronidases, is observed (Allingham et al., 2006). This is indicative

that hyaluronan levels will be increased in proliferating cells and decreased when cells are static. The localization of the hyaluronan is visible using fluorescent hyaluronan and can be seen to create a pericellular matrix composed of a tangled mass of fibres (Evanko and Wight, 1999). This fluctuation of hyaluronan concentration within cells is seen during the cell cycle with the concentration peaking most during mitosis (Toole, 2001). This pericellular matrix will create a highly hydrated environment due to hyaluronan having a net negative charge causing destabilisation of cellular attachment, contributing to cell rounding and providing an environment for the presentation of growth and differentiation factors (Toole, 2001, Evanko et al., 1999, Satoh et al., 2004). The importance of hyaluronan for cell proliferation has also been demonstrated. When the cell cycle is interrupted, the concentration of hyaluronan drops as it is no longer required for mitosis. Likewise if hyaluronan synthesis is stopped by inhibition of hyaluronan synthase the cell cycle is terminated showing a direct link between the production of hyaluronan and cell division (Brecht et al., 1986).

1.2.1.2.2 Cancer

It is well documented that hyaluronan levels are elevated in many forms of cancerous tumour including prostate, breast, bladder and ovary. Sustained increased expression of hyaluronan is also closely associated with the progression and metastasis of tumours. It is believed that hyaluronan serves several functions in tumour development. In the extracellular matrix, hyaluronan is believed to provide a structural scaffold on which the tumour cells can adhere and proliferate. The extracellular matrix also provides the developing tumour with a source of growth factors and cytokines needed for cell development (Paiva et al., 2005). The scaffold of hyaluronan in the extracellular matrix also creates an ideal framework for migration for tumour cells by its ability to promote receptor mediated migration (Wein et al., 2006). Hyaluronan also provides cancerous cells with an immunological 'coat'. The hyaluronan can cover the outside of the cell either by binding to receptors on the cell surface or newly synthesised hyaluronan which is being extruded from the plasma membrane can remain attached. As the hyaluronan is biologically normal unlike the cell which it

coats, it is viewed as immunologically safe and the cell goes undetected by the immune system (Wein et al., 2006, Heldin, 2003). Another role of hyaluronan in tumour development is that of angiogenesis. Low molecular weight fragments of hyaluronan have been shown to stimulate angiogenesis, a crucial step in the development of a blood supply to a tumour (Paiva et al., 2005, Tawada et al., 2002, Mahoney et al., 2001). The low weight fragments are seen due to an increase in the expression of the hyaluronidase genes which allow for the degradation of hyaluronan into these angiogenic fragments (Lokeshwar et al., 1999).

1.2.1.2.3 Wound healing

Small and large molecular weight hyaluronan molecules both have roles in wound healing. During initial injury high molecular weight hyaluronan levels increase sharply and serves the purpose of binding fibrinogen for clot formation. Following this initial stage of wound healing the levels of low molecular weight hyaluronan begins to increase and the inflammatory process and angiogenesis take place creating new blood supply for the new tissue formation (Frankova et al., 2006, Stern et al., 2006).

1.2.1.2.4 Inflammation

At sites of inflammation it has been observed that the concentration of hyaluronan rises (Termeer et al., 2003, Jentsch et al., 2003). This increase in the level of hyaluronan is also associated with the formation of hyaluronan cables created by the cross-linking of many fibres. These cables are believed to behave as a structural platform for inflammatory cells and plays a role in the trafficking of such cells (Day and de la Motte, 2005, Mummert, 2005). The purpose of this platform is not entirely clear. It is apparent that leukocytes needed for inflammation will become bound to this network of hyaluronan via the CD44 receptors on their surface. Two theories have been put forward to attempt to explain the purpose of the bound leukocytes. One possibility is that the hyaluronan acts to mediate the inflammatory response by restricting the

leukocytes access to the damaged area and this would also serve the purpose of protecting the extracellular matrix from damage by inflammatory cells (Day and de la Motte, 2005). The alternative is that the hyaluronan cables provide an anchorage point for the leukocytes to mediate their migration towards the site of damage (Hascall et al., 2004, Pure and Cuff, 2001).

1.2.2 Chondroitin

Chondroitin is another class of GAG and is composed of repeating disaccharide subunits of glucuronic acid and *N* acetyl-galactosamine linked by a β1-4 linkage. The *N*-acetyl galactosamine can be sulphated in one of two positions, either on the 4^{th} carbon or the 6^{th} carbon yielding chondroitin-4-sulphate (chondroitin sulphate A) and chondroitin-6-sulphate (chondrotin sulphate C) respectively or can be sulphated in both positions (chondroitin sulphate E) (Figure 1.3)

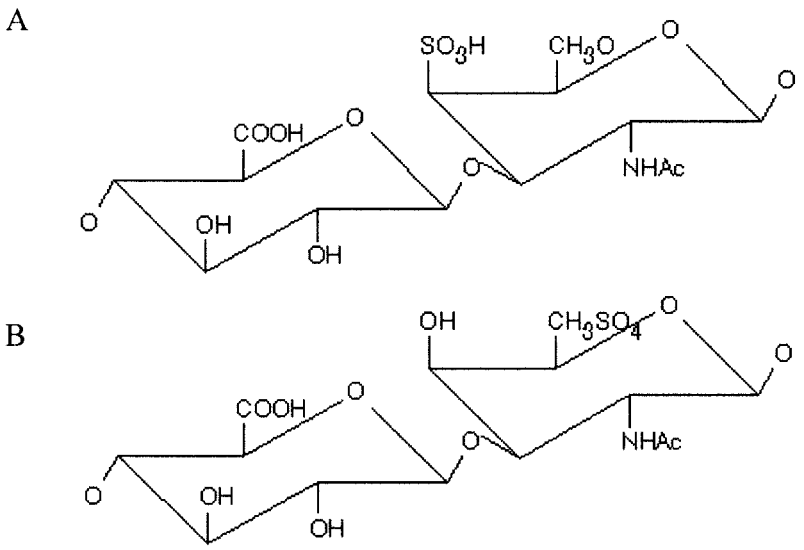


Figure 1.3: The structure of A: Chondroitin 4 sulphate and B: Chondroitin 6 sulphate. Taken from http://biol.lancs.ac.uk/gig/pages/gag/cspages.htm

The sulphate group along with the carboxylate group of the glucuronic acid contributes to the negative charge of the molecule allowing for the retention of

water in the molecule, similar to hyaluronan. The sulphation pattern of chondroitin varies depending on the age and the location of the chondroitin. Mature chondroitin presents with a marked increase in levels of sulphation at the sixth carbon position. The position of the chondrotin in a sample appears to influence the sulphation with deep layers being up to four times more sulphated than that found in the upper layers of a sample. It has also been noted that the extracellular matrix of compact bones (such as those following fracture) is rich in chondroitin-4-sulphate (Bali et al., 2001).

Chondroitin is synthesized by the action of glycosyltransferases and sulfotransferases which combine UDP derivatives of glucuronic acid and N acetyl galactosamine to form the saccharide polymer. In cartilage, and the extracellular matrix, approximately 100 chains each comprised of between 50 and 60 disaccharide units covalently attach to a polypeptide backbone of approximately 2000 amino acid residues. The attachment occurs between either the xylose or the galactose residue of the polysaccharide and a serine or a threonine of the core protein. This macromolecule is known as aggrecan and as previously mentioned attaches to hyaluronan via a link protein (see section 1.2.1.1.1) . This link protein is found at the end of the core protein and forms a globular region which will stretch around and accommodate 5 disaccharide units of hyaluronan. Approximately 100 of these core proteins will attach to a hyaluronan chain given a molecule with an average molecular weight of 2×10^6 Daltons (Bali et al., 2001). The core protein is synthesized in membrane bound ribosomes and passed through into the lumen of the endoplasmic reticulum. The chondroitin sulphate chains are assembled in the Golgi apparatus and the two combined. Aggrecan is found in the extracellular matrix particularly in cartilage where it provides load bearing properties. It is also found on cell surfaces and in neural tissue. Its presence has been shown to provide similar cell proliferation effects as hyaluronan in that it appears to stimulate neural outgrowth (Lauder et al., 2000). In the extracellular matrix, chondroitin sulphate does not have the same high levels of viscosity as seen in hyaluronan suggesting that it is not involved in the modulation of tissue and fluid viscosity by the attraction of water like hyaluronan but it has been observed that chondroitin increases the viscosity of hyaluronan whilst remaining at a low viscosity itself. This is believed to be as

a result of the interaction of hydrophobic patches on both the hyaluronan and the chondroitin creating a more viscous solution when combined compared to the two when observed alone (Nishimura et al., 1998). This is indicative of the complex nature of the extracellular matrix and the interactions between molecule creating the unique properties that it has.

1.3 GAG-degrading enzymes can be found in the carbohydrate-active enzymes classification

This system (http://afmb.cnrs-mrs.fr/CAZY/) categorises putative enzymes and known enzymes into classes and families based on amino acid sequence similarities (Henrissat, 1991; Henrissat *et al.*, 1998; Coutinho and Henrissat, 1999). The degree of similarity is determined by gapped BLAST and psi BLAST searches against a library of protein modules with known activity and also the non-redundant protein database of the National Centre for Biotechnology Information (NCBI). Complementary sequences (those with an E value below 0.001 indicating that the probability of these being different is less then 0.1%) are deemed to be the same module and categorised as such. In the instance of low similarity results, hydrophobic cluster analysis is also performed to detect structural similarities where the primary sequence similarity is low (Coutinho and Henrissat, 1999). The classes formed as a result of this classification are glycoside hydrolases (GHs), polysaccharide lyases (PLs), glycosyltransferases (GTs), carbohydrate esterases (CEs) and carbohydrate binding modules (CBMs). These classes group the enzymes loosely into predicted activity. Within the classes there are also families formed, the members of which share primary sequence homology and hence the same overall three-dimensional fold (Coutinho and Henrissat, 1999). This method of classification is different compared to that of the traditional EC classification system which groups enzymes based primarily on substrate specificity and has many advantages over the EC method. One such major advantage of classification based on sequence similarities is that it allows for a grouping of enzymes with different EC numbers (and therefore different substrate specificities) into polyspecific families (similar enzymes but with different substrate specificities) (Henrissat, 1991). This can

facilitate insights into the divergent evolution of enzyme families. Enzymes within a family may have very different substrate specificities but be closely enough related to be considered divergent of one historical ancestor. It can also highlight convergent evolution whereby enzymes with very similar function and substrate specificity which would normally be grouped together by the EC classification system are shown to belong to several distinctly different families. Underpinning this classification system is the prediction that sequence determines structure. Therefore from this amino acid sequence, key information regarding the structure and mechanism of unknown enzymes could be elucidated (Henrissat and Davies, 1997). This method of classification means that all members of a family must share common properties and that predictions can be made of newly classified members if the properties of one or more of the family are determined.

The efficient and complete degradation of polysaccharides, whether they are from plant or animal cells, requires organisms to produce an extensive array of endo- or exo-acting enzymes. Degrading carbohydrate active enzymes can be described as being either endo- or exo- in their mode of action. Endo-acting enzymes cleave the substrate in random internal positions yielding a variety of differently sized oligosacchraides. Exo-acting enzymes will cleave and remove one or more residues from either the reducing or non-reducing end of a polysaccharide. This will yield one size oligosaccharide which will increase in concentration over time.

1.3.1 GAG-degrading polysaccharide lyases

PLs in general cleave the glycosidic bond found in polysaccharides between adjacent disaccharide units via a β-elimination mechanism which results in the formation of a double bond between the 4^{th} and 5^{th} carbon of the uronic acid residue at what is now the newly formed non-reducing end (Baker et al., 2002, Linker et al., 1960). To date (September 2007) there are eighteen families based on sequence similarities. In comparison to the number of families in the other classes (with the exception of the CE's); GH:110, GT:90 and CBM:49, this

number of families is relatively small. The limited numbers of PLs can be explained by the rigid requirements for the substrates. They require the substrate to have free or methyl esterified C5 carboxylate groups and this limits the range of substrates available to them hence leading to a lower ubiquity in nature when compared to the other carbohydrate active enzymes.

The substrate specificity of GAG-degrading polysaccharide lyases is not rigid. There is a degree of flexibility amongst the substrates upon which they will work. For example some hyaluronate lyases, as well as degrading hyaluronan, also demonstrate activity against chondroitin sulphate due to its structure of repeating disaccharide units (D-glucuronic acid-β1,3-*N*-acetyl-galactosamine) being similar to that of hyaluronan (D-glucuronic acid-β1,3-*N*-acetyl-glucosamine) with the difference being in the anomeric configuration of the C4 of the glycosamine. Chondroitin often takes a sulphated form at either the fourth, sixth or both carbons of *N*-acetyl galactosamine. Some hyaluronate lyases are able to degrade all three of these arrangements by an endolytic mechanism followed by processive cleavage (Baker et al., 2002, Rigden and Jedrzejas, 2003b).

1.3.1.1 Architecture of GAG-degrading polysaccharide lyase family 8 (PL8)

The PLs are most often comprised of a multi domain structure. The general structure consists of three domains joined by linker sequences. There is commonly a catalytic domain joined to 2 ancillary domains believed to be involved in part in substrate binding. Catalytic domains have been shown to be comprised entirely of either α helices (Rigden and Jedrzejas, 2003a) or β sheets (McDonough et al., 2004) or a mixture of both.

The structures of the *Streptococcus pneumoniae* and *Streptococcus agalactiae* hyaluronate lyases have shown that they too possess the three domain structure previously mention with a catalytic all α helix domain flanked on either side with an all β sheet domain (one large and 1 small). The α helices of the α domain are found to be in a twisted α/α barrel which is incompletely wound at one end

making it slightly wider than the other. The outer layer is unidirectional whilst the inner layer helices run opposite to one another.

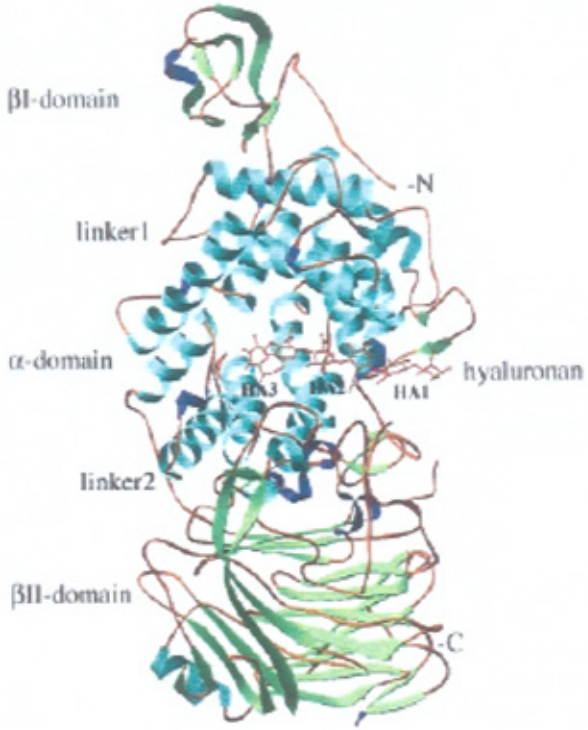


Figure 1.4: Ribbon diagram of the structure of *S. agalactiae* hyaluronate lyase (Taken from Mello *et al.*, 2002)

1.3.1.1.1.1 Architecture of the active site of PL8s

Within the cleft active site of hyaluronate lyases there are three distinct regions. Three aromatic residues are found opposite three catalytic residues with three negatively charged residues found away from the others forming a negative patch. The three aromatic residues; 2 tryptophans and 1 phenylalanine form a hydrophobic patch opposite the catalytic residues. This region interacts with the hydrophobic patch found in hyaluronan and allows for precise positioning of the substrate within the cleft. The three catalytic residues; asparagine, histidine and tyrosine are all positively charged and contribute to the attraction and binding of the negatively charged hyaluronan (Li and Jedrzejas, 2001). They are also involved in the method of catalysis which is via proton donation and acceptance. The negative patch formed by residues glutamine, aspartic acid and threonine is

found at the end of the cleft in the vicinity of the reducing end of the polysaccharide. Due to the net negative charge of the hyaluronan repulsion occurs and the role of this patch is believed to be involved in product release following catalysis (Mello et al., 2002, Rigden and Jedrzejas, 2003a).

1.3.1.2 Hyaluronate lyase mechanism of action

The mechanism of action of hyaluronate lyases was first proposed to be successive proton donation and acceptance by a single histidine residue (Greiling et al., 1975) but a more accepted method now is of a procedure involving not only a single histidine which acts as a general base (Lin et al., 1997) but also a tyrosine and an asparagine residue (Ponnuraj and Jedrzejas, 2000) (Figure 1.5).

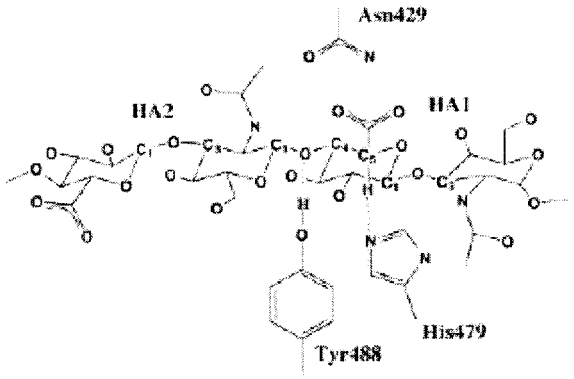


Figure 1.5: Schematic to show the degradation of hyaluronan by *S. agalactiae* hyaluronan lyase SagHL taken from Li & Jedrzejas, 2001.

The mechanism can be broken down into 5 successive steps. Step 1 involves the attraction of negatively charged hyaluronan into the cleft by the positively charged residues located within it. The second step sees the aromatic patch not only securely attach the hyaluronan molecule into the cleft but also orientate it in to the correct degradation position. In this correct position the carboxylate group of the glucuronate group is ideally positioned next to the catalytic asparagine residue, the glycosidic oxygen is near to the tyrosine and the C5 of the glucuronate is in proximity to the histidine. The side chain of the asparagine forms a hydrogen bond with the carboxylate group of the glucuronate attracting

the negative charges of the carboxylate away from the C5. This results in the formation of an acidic C5 hydrogen. The third step is the removal of a proton from the C5 by histidine. Simultaneously to this in step 4 the tyrosine residue interacting with the oxygen of the glycosidic bond donates a proton to the glycosidic oxygen thereby breaking the glycosidic bond. These steps leave both the C4 and the C5 of the glucuronic acid molecule in a proton depleted state which then transforms into a carbon-carbon double bond. The fifth and final step in the degradation involves the return of the enzyme to its original state. This involves the histidine losing the abstracted proton and the tyrosine attracting a proton (Li and Jedrzejas, 2001, Li et al., 2000).

1.3.2 GAG-degrading glycoside hydrolases

Glycoside hydrolases (GHs) are a class of enzymes that are responsible for the hydrolysis of glycosidic bonds in carbohydrates. They are involved in the metabolism of plant and animal polysaccharides for use as a carbon and energy source within the host organism. Due to a wide and varied substrate collection there is a large assortment of GHs responsible for the hydrolysis of the glycosidic bonds. To date (September 2007) there are 110 families of GHs which are further grouped into 'clans'. These clans are created as a result of the determination of the 3D structure for 46 families. Many of these family structures showed a distinct similarity to others from different families and hence have been assigned to clans or superfamilies. The conserved structures within these clans is indicative that the members have divergently evolved and that although the main structures of the enzymes are different the regions under selective pressure, ie the active site regions have maintained integrity (Bourne and Henrissat, 2001, Henrissat and Davies, 1997).

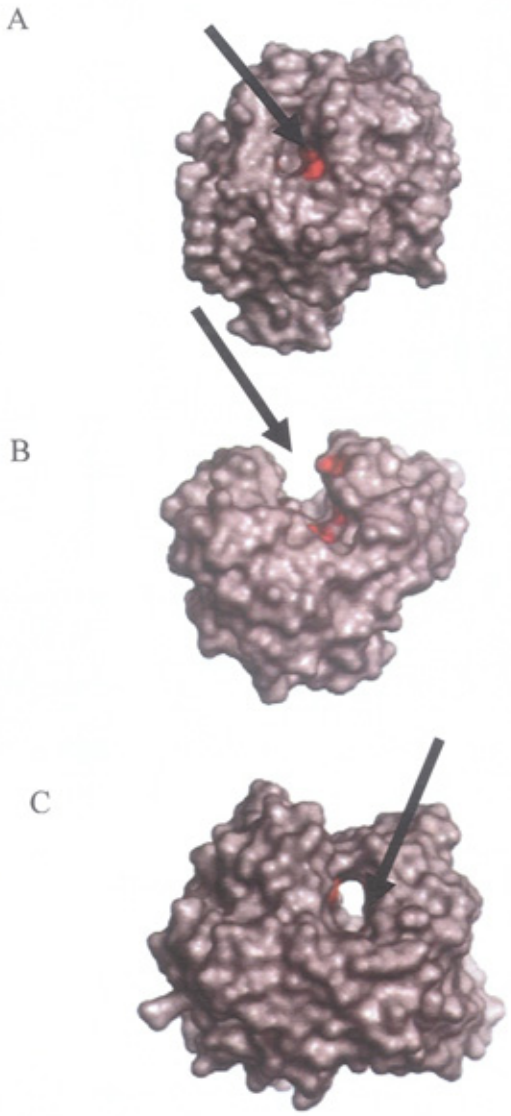


Figure 1.6: Active sites of glycosyl hydrolases. A: pocket (glucoamylase from *Aspergillus awamori*); B: cleft (endo-1,4-β-D-glucanase from *Thermomonospora fusca*); and C: tunnel (cellobiohydrolase from *Trichoderma reesei*). Active site (indicated by arrows) topologies displayed. Taken from Davies and Henrissat (1995)

1.3.2.1 Active site topologies of GHs

The 3D structure data for GHs have revealed three general active site topologies: pocket, cleft (or groove) and tunnel. These active sites are found in the different enzymes and reflect their subsequent digestion products: disaccharidases, endo-acting and exo-acting enzymes respectively (Figure 1.6). The pocket topology is perfectly designed to recognise and bind just the chain ends of a polysaccharide. Digestion following this yields only disaccharides. The cleft structure has an open active site with little restriction on the binding of substrate. This structure is typical of the endo-acting enzymes whereby the polysaccharide chains bind randomly within the cleft, (although they will be in steric arrangement to match the catalytic residues) and digestion from this results in the formation of variable length oligosaccharides. The tunnel structure is found in exo-acting enzymes and is similar in structure to the cleft structure but it has additional polypeptide loops which form a top to the cleft giving the characteristic tunnel formation. The active site within this structure allows the entry of a glycan chain at one end which progresses through the tunnel and releases digested products from the end (Davies and Henrissat, 1995).

1.3.2.2 Catalytic mechanisms

Most of the enzymes in the 108 families of the GHs class have been demonstrated to catalyse the glycosidic bond with either retention or inversion of the configuration of the C1 of the bond being broken. Cleavage occurs via a general acid/base assisted catalysis method (Vasella et al., 2002, Warren, 1996). Figure 1.6 depicts both of these mechanisms. The two mechanisms both require two carboxylic acid residues in the active site of the enzyme for catalysis. However, the catalytic roles of these two residues vary depending on the mechanism involved. The majority of GHs studied have shown these catalytic amino acids are only ever glutamic acid and/or aspartic acid (Ozimek et al., 2006, Nakai et al., 2005).

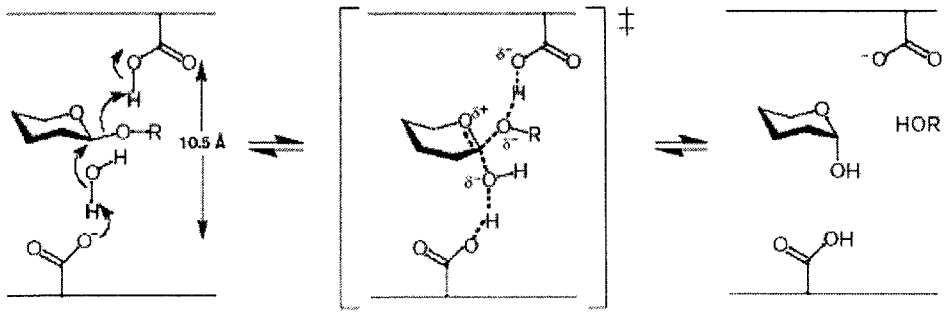
1.3.2.2.1 Inverting mechanism

Inverting enzymes catalyse the hydrolysis of the glycosidic bond by a single displacement mechanism resulting in overall inversion of the configuration. The two catalytic residues of the inverting enzyme are positioned ~ 10 Å apart which allows for the inclusion of a water molecule between the sugar and the general base needed for catalysis. One of the carboxylic acid residues acts as a general base whilst the other acts as a general acid. The process of cleavage is a single step where the general acid protonates the glycosidic oxygen and simultaneously to this the general base causes the formation of a hydroxyl group from the water molecule associated in the active site. This hydroxyl group then performs a nucleophilic attack on the carbon of the uronic acid bringing about the cleavage of the glycosidic bond leaving a product with the opposite stereochemistry (Rye and Withers, 2000).

1.3.2.2.2 Retaining mechanism

Retaining enzymes catalyse the cleavage of glycosidic bonds via a double-displacement mechanism, which leads to the overall retention of the configuration. In comparison to the inverting enzymes, the two catalytic residues in the active site of a retaining enzyme are positioned ~ 5.5 Å apart. This narrower size is due to the fact that no water is needed between the sugar and the catalytic residue and also ensures the close proximity between the nucleophilic residue and the sugar needed for the cleavage. One of the two key carboxylic acid residues functions as both a general acid and a general base, while the other acts as a nucleophile (Withers, 2001). The mechanism for retaining enzymes involves a two-step process. In the first step, the acid/base catalytic residue acts as an acid and causes protonation of the glycosidic oxygen, resulting in bond cleavage. Simultaneously to this, the nucleophilic residue acts upon the anomeric carbon of the glycosidic bond forming a covalent linkage between the sugar and the enzyme (Davies et al., 1998). The acid/base catalytic residue then functions as a general base and abstracts a proton from a water molecule, which acts upon the sugar-enzyme intermediate. This cleavage results in the generation of a sugar product with the same configuration as the substrate.

A.



B.

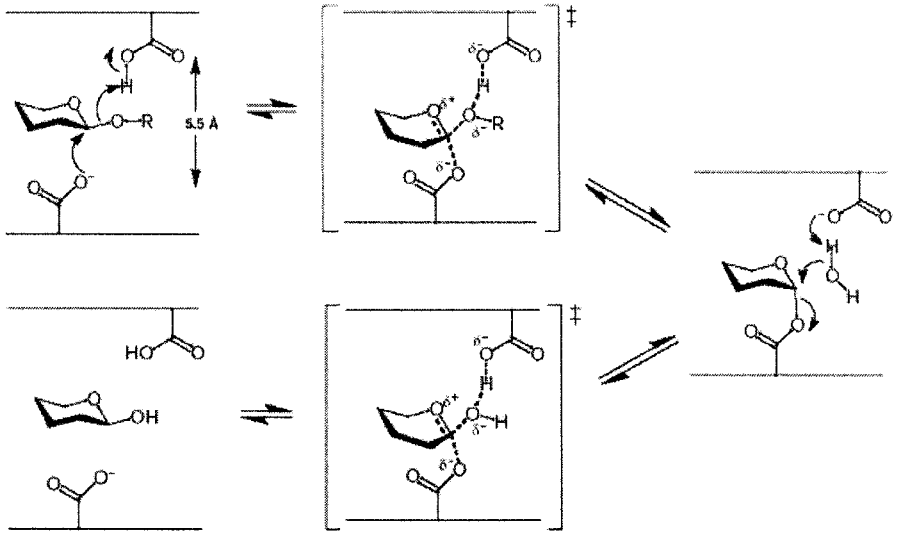


Figure 1.7: The mechanisms of glycoside hydrolases. A: inverting mechanism B: retaining mechanism.
Taken from http://www.cazy.org/fam/ghf_INV_RET.html#inv

1.3.2.3 Glycoside hydrolases of glycosaminoglycans

A wide array of bacteria, bacteriophage, fungi, invertebrates and vertebrates synthesise numerous enzymes that hydrolyse GAGs reflective of the wide distribution of the substrates in nature GAG-degrading GHs are found in CAZy famlies 56, 79 and 84. The degradation of GAGs, in particular hyaluronan, by hydrolases and lyases may contribute towards the overall virulence of the pathogenic organisms.

1.3.2.3.1 Hyaluronidases

The term hyaluronidase was introduced to denote enzymes that degrade hyaluronan (Frost et al., 1996). Hyaluronidases were first identified as a 'spreading factor' in an extract from mammalian testes that facilitated diffusion of subcutaneously injected dyes. It became apparent thereafter that the spreading factor was an enzyme and that its substrate was hyaluronan (Frost et al., 1996). Hyaluronidases are produced by groups A, B, C and G streptococci, staphylococci, candida and clostridia, plus many others (Hynes and Walton, 2000).

1.3.2.3.1.1 Invertebrate hyaluronidases

Hyaluronidases have been found in a wide variety of invetebrates; bees, wasps and hornets, leeches and hookworm (Hotez et al., 1992). The hyaluronidases of bees, wasps and hornets are found in the venom whereas the hyaluronidase of leech and hookworm are secreted from the organism. In all cases the hyaluronidase is believed to be a spreading or dissemination factor. In the case of the venom this spreading would allow for the systemic spread of other venom components throughout the body (Gmachl and Kreil, 1993). In the case of the leech and the hookworm the secretion of hyaluronidases allow for either the improved extraction of blood through the skin or to allow for access through the skin respectively.

1.3.2.3.1.2 Vertebrate hyaluronidases

Hyaluronidases are found in nearly all snake venoms, but only a few have been characterised. Snake venom hyaluronidases are by themselves non toxic and have been shown to act as spreading factors in the same way as seen in bee venom hyaluronidase and serving the same purpose of venom dissemination (Girish et al., 2002).

Hyaluronidases have been found in a vast range of sites in mammals. Protein PH-20 (GH56) which shows homology to the honey bee venom hyaluronidase (Gmachl and Kreil, 1993), is present on the plasma membrane of sperm, and this protein was shown to be a hyaluronidase. The oocyte of mammals is surrounded by a cumulus extracellular matrix that extends past the zona pellucida. The cumulus ECM is rich in hyaluronan and the sperm must penetrate this matrix to gain entry to the oocyte. The PH-20 hyaluronidase of sperm allows for this penetration (Cherr et al., 2001, Arming et al., 1997, Zheng et al., 2001).

There are six hyaluronidase like genes in the human and mouse genomes HYAL1-4, PH-20/SPAM1 and an expressed pseudo gene HYALP1. HYAL1 and hyaluronan turnover by it has been shown to promote tumour cell cycling and as such may be involved in the malignancy of tumours (Lin and Stern, 2001). HYAL2 is less active than the others but still shows an important role in tumour malignancy (Lepperdinger et al., 2001). HYAL3 is believed to be involved in stem cell regulation due to high

level expression in testes and bone marrow (Csoka et al., 2001). HYAL4 is classed as a hyaluronidase but evidence suggests that it is substrate limited to just chondroitin (Csoka et al., 2001).

1.3.2.3.2 Unsaturated glucuronyl hydrolase (UGL)

An unusual group of enzymes found within the class of glycoside hydrolase and only in families 88 and 105 are the unsaturated glucuronyl hydrolases (UGL). These enzymes cleave the glycosidic bond only between an unsaturated uronic acid formed by polysaccharide lyase degradation and its neighbouring sugar. This removes the unsaturated residue from the polysaccharide leaving the polysaccharide ready for further degradation. This enzyme is intracellular and allows for the complete depolymerisation of polysaccharides which allows for better utilisation as a carbon source (Hashimoto et al., 1999b, Nankai et al., 1999). Its activity is usually seen on tetrasaccharides of gellan, xanthan, hyaluronan and chondroitin but has also been observed on disaccharides of hyaluronan and chondroitin (Hashimoto et al., 1999a). As the UGL acts on the unsaturated products of gellan, xanthan, hyaluronan and chondroitin it is likely that they will be produced by organisms capable of producing polysaccharide lyases and are used as a route for complete depolymerisation and assimilation of the monosaccharides. They have also been observed in strains of streptococcus which are capable of causing severe infection so not only are they useful for depolymerisation for carbon utilisation but may also aid in the degradation of ECM components and facilitate the spread of the bacteria (Mori et al., 2003).

1.3.2.3.2.1 Structure of UGL and its active site architecture

UGL from *Bacillus* spp have been identified and their structures elucidated. The enzymes have an α_6/α_6 barrel formation with the 6 outer and 6 inner helices running unidirectional but opposite to one another. These helices are linked in a nearest neighbour formation by both long and short loops with the long loops on one side of the barrel and the short loops on the other side. The long loops form a deep conical shaped pocket which is the location of the active site (Itoh et al., 2006b, Itoh et al., 2004). Unlike other glycoside hydrolases there is no direct interaction between the glycosidic bond and the active site residues, the sugar is simply positioned within the active site by interactions between the carboxyl group of the uronic acid and the uronic acid ring (Itoh et al., 2006a). The inside of the active site is positively charged and hence will attract negatively charged glycosaminoglycans into it. The carboxyl group of the uronic acid is positioned ideally to the positive side chain of an arginine residue which allows for the creation of a salt bridge (Itoh et al., 2006c). The carboxyl group of the uronic acid forms multiple hydrogen bonds with residues within the active site thereby stabilising the sugar in place (Itoh et al., 2006a). The sugar is further stabilised in place by a stacked tyrosine in the active site which binds to the sugar ring for improved stability (Itoh et al., 2006c). Following cleavage the pyranose ring released spontaneously converts to a chain form due to the instability caused by the keto-enol equilibrium (Itoh et al., 2006a). This makes understanding of the anomeric configuration and hence the mechanisms of the cleavage very difficult. It is impossible to see the monosaccharide as a ring within the active site and understand the steps involved in cleavage so as yet the mechanism is elusive. The enzyme itself is predicted to have an inverting mechanism. This is due to the distance between the possible catalytic groups of the known enzymes being positioned approximately 7.3 Å apart which is similar to that seen in other inverting α_6/α_6 barrel topology enzymes (Itoh et al., 2004).

1.4 Streptococci

Streptococci are Gram-positive catalase negative facultative anaerobes (Perera and Hay, 2005). Streptococcal species are primarily classified by their haemolytic properties. There are 3 main haemolytic groups; α which reduce the iron in haemoglobin, β which rupture red blood cells completely and γ which have no effect on blood cells. They can also be further classified by the Lancefield serological grouping method. Developed by Rebecca Lancefield in 1962 this classification is based on the specific carbohydrates on the bacterial cell wall and constitutes groups A-T. Streptococci which do not react with the Lancefield antibodies include *Streptococcus mutans*, which has been shown to be the primary aetiological cause of oral caries (Shah and Russell, 2004). Streptococcal species are normal commensal flora of the skin, intestines, mouth and upper respiratory tract. However, some strains can become pathogenic leading to a wide variety of disease manifestation.

Primary infections	Impetigo Erysipelas Cellulitis Vulvovaginitis Ulcers Necrotizing fasciitis
Secondary infections	Eczema Infestations Ulcers
Toxin induced damage	Scarlet fever Toxic shock-like syndrome
Allergic hypersensitivity	Vasculitis
Contribution to dermatosis	Psoriasis

Table 1.1: Table showing the diseases associated with human streptococci (Adapted from Perera & Hay, 2005)

The 2 primary Lancefield groups looked at in this study are group A (*Streptococcus pyogenes*) and group C (*Streptococcus equi*).

1.4.1 *Streptococcus pyogenes*

S. pyogenes is a prevalent pathogen which is estimated to cause severe infections in more than 18 million people per year with more than 0.5 million deaths (Zhang et al., 2007) The infections it causes can be grouped into 4 distinct categories; superficial, deep, toxin-mediated and immunologically mediated. The superficial less severe diseases include impetigo, pharyngitis and erysipelas. Deep infections include cellulitis, necrotizing fasciitis and meningitis. Toxin mediated infections are scarlet fever and toxic shock-like syndrome. The immunologically mediated diseases include rheumatic fever (Efstratiou, 2000). It causes a range of diseases from the less severe; impetigo and pharyngitis, to the severe cases of acute cellulitis and deep tissue infections. Fatal infections can also be caused by *S. pyogenes* including streptococcal toxic shock (Cohen et al., 2005) and necrotizing fasciitis which account for a mortality of 10-13 % of *S. pyogenes* infections (Miller and Neely, 2004). *S. pyogenes* are known to produce a wide range of extracellular proteins which are known to increase the virulence of the organism and account for the wide degree of variability in pathogenicity between strain. Approximately 10 % of the genome of *S. pyogenes* is composed of bacteriophage genetic material in the form of 4 complete or partial bacteriophage genomes (Ferretti et al., 2001).

1.4.2 *Streptococcus equi*

S. equi is a Gram-positive β haemolytic group C streptococcus. It is an obligate pathogen of equids which can be carried by approximately 10 % of horses following an infection for several months with no disease manifestations. The horse can then shed and pass on the bacteria within stables propagating further infections of the subsequent disease 'strangles' (Timoney, 1988, Kelly et al., 2006, Flock et al., 2004). Strangles manifests itself as a nasopharyngeal infection which infiltrates the lymph nodes of the head and neck. Subsequent bacterial multiplication causes abscess formation which, after approximately 2-3 weeks will rupture and drain through sinus tracts either to the upper respiratory mucosa or through the skin (Flanagan et al., 1998). It is these large neck abscesses which contribute to the name of the disease as they quite often cause obstruction of the upper airway leading to a strangled breathing sound. In approximately 10 % of infections various complications can

occur. Secondary abscesses formation on other body organs which rupture causing death is a complication known as 'bastard strangles' (Chanter et al., 1995). Petechial haemorrhaging and oedema of the limbs and gums can occur which is associated with the production of antibody-M-like protein complexes. These then block the capillary beds causing further oedema, circulatory failure and ultimately death (Harrington et al., 2002).

S. equi is sensitive to penicillin and other antibiotics but treatment with such is usually ineffective with the disease progression only abating or being delayed until the treatment is removed. This is probably due to the lack of adequate vascularity within the abscess so the antibiotic cannot penetrate deep enough to reach therapeutic levels (Harrington et al., 2002).

1.4.2.1 Other diseases associated with *S. equi*

S. equi is normally only ever associated with causing strangles in equids but recent reports suggest that it may also be capable of infecting other hosts and causing different diseases. There has been a reported case of strangles in a canine host suggestive of a new host for what was previously contained as an equine pathogen (Ladlow et al., 2006). The pathogen has also been seen to cause streptococcal meningitis in humans believed to have been contracted from contact with a carrier horse (Elsayed et al., 2003).

1.4.3 Streptococcal virulence factors

Streptococcal species as discussed are capable of causing a wide variety of pathologies ranging, in the case of *S. pyogenes*, from the superficial such as throat infections, to the life threatening, such as necrotizing fasciitis. Differences between related species and within a species in the resulting pathology of infection are often attributed to virulence factors. These factors can be either chromosomally encoded or in some cases bacteriophage encoded (see section 1.5.4). A list of the principal streptococcal virulence factors is given in Table 1.2.

Virulence factor	Function
M and M like proteins	Adhesion, internalisation, antiphagocytic
Hyaluronan capsule	Antiphagocytic, adhesion
Fibronectin binding protein	Adhesion, internalisation
C5a peptidase	Antiphagocytic
Hyaluronan degrading enzymes	Dissemination
Streptococcal pyrogenic exotoxins	Systemic toxicity
Streptolysin	Systemic toxicity

Table 1.2: Streptococcal virulence factors and their proposed pathogenic functions (Medina, 2004).

1.4.3.1 M and M-like proteins

The M protein is an antiphagocytic cell surface molecule which shows a high degree of variation between species. The M protein was identified in the 1920s and was the basis of Rebecca Lancefield's work on which the Lancefield classification system is based. Further to this the high degree of variation allows for further classification of streptococcal strains. For example, within GAS, more than 130 different M types have been identified based on variations in the M protein on the surface of the bacteria. It has also been noted that there is a non-random association of different M serotypes and the infection type caused. Serotype M1 is a common cause of pharyngitis and invasive infections while M18 strains are more often associated with acute rheumatic fever outbreaks. M3 strains are commonly seen in unusually severe infections with high mortality rates (Beres et al., 2002).

The structure of the M protein is of a coiled coil comprised of a hyper variable region (HVR) at the N terminal followed by three repeat regions designated A, B and C-repeats. These repeats are then followed by a cell wall anchor motif and a stretch of hydrophobic amino acids thought to be embedded in the cell membrane (Vohra et al., 2005, Courtney et al., 2006). The M protein within a strain is stable but the HVR differs between strains and determines the strain type. Antibodies directed against the HVR are protective and the immunity is type specific. This antigenic variation has thus far hindered the development of an effective vaccine (Areschoug et al., 2004). The conserved repeat regions are another possible target for vaccine development. There have been 2 opsonic peptide sequences identified in the C-repeat region of the structure which when administered can cause the production of antibodies which are protective and are capable of protecting against a variety of GAS M types (Vohra et al., 2005).

The M protein of *S. equi* occurs in only one form which should make the development of a vaccine easier due to the lack of hypervariability but *S. zooepidemicus* demonstrates a similar heterogeneity of M protein types as seen in *S. pyogenes* leading to the same vaccine development problems. Although vaccine development against *S. equi* is straightforward and vaccines are effective in the stimulation of antibodies, field exposure has given poor results. Vaccination results in only a two-fold reduction in outbreak. This is thought to be due to an

intramuscular or subcutaneous delivery, being unable to stimulate nasopharyngeal antibodies which are needed at this site of infection (Timoney and Mukhtar, 1993).

The importance of the M protein in the development of *S. equi* infection is apparent in the fact that some carriers of *S. equi* demonstrate clones which have truncated M proteins which when studied tend not to be the causative agent of strangle outbreaks. This deletion occurs at the N terminus and causes a reduction of approximately 20 amino acids further confirming the importance of this N terminus region in the infection mechanism (Chanter et al., 2000). Mechanistically the inhibition of phagocytosis by M proteins is poorly understood primarily due to the hyper variability of the HVR and its role in the inhibition. It has been found that many HVRs are able to specifically bind to the same ligand despite their variability. This ligand is human C4b binding protein (C4BP) which is a plasma protein which inhibits the activation and progression of the classical complement pathway and bacteria bound C4BP retains this function causing a regional inhibition of complement activation (Areschoug et al., 2004).

As well as being of importance as an antiphagocytic agent the M protein has also been demonstrated to be involved in adherence and internalization of bacterial cells although not all of the M protein types behave in the same way in regards to these functions. M1 and M6 are reported to mediate adherence and internalization whilst M49 and M18 neither hinder nor mediate such processes (Eyal et al., 2003).

1.4.3.2 Hyaluronan capsule

GAS and GCS are both encapsulated bacteria (Anzai et al., 1999b). The capsule is composed of hyaluronan which is chemically similar to that found in human connective tissue. The production of this capsule is controlled by three different genes on an operon. The three genes are termed *hasA*, *hasB* and *hasC*. *HasA* encodes a hyaluronate synthase, *hasB* a UDP-glucose dehydrogenase and *hasC* encodes a UDP-glucose pyrophosphorylase (Bisno et al., 2003). Structurally the hyaluronan of the capsule assumes a tape like two fold helix with the hydrophobic groups located on the upper and lower surfaces of the tape and the polar groups arranged along the tape edges. The helices then stack in tertiary structures which can resemble the antiparallel β sheets of polypeptides (Scott et al., 2003). Hyaluronan produced by the

cells is exported from the cell to form the capsule by an ABC transporter protein which is encoded upstream of the *has* operon (Ouskova et al., 2004). Different streptococcal strains vary greatly in their degree of encapsulation with the highest amount of encapsulation producing a mucoidal appearance to the bacteria when grown on solid media. The role of the capsule in the evasion of host defences as an immunological coat has been well recognised but only recently has it been fully demonstrated using modern molecular biology. A comparison between encapsulated and non-encapsulated strains has shown its importance as the acapsular strain lost all ability to resist phagocytosis and showed a 100-fold reduction in virulence (Bisno et al., 2003).

As well as serving an antiphagocytic function it has also been demonstrated that the hyaluronan capsule may serve a purpose in adherence. It was noted that the hyaluronan capsule inhibited the adhesion properties of the M proteins and other adhesins of highly encapsulated bacteria. This fails to conform to the understanding that highly encapsulated strains are highly pathogenic as adhesion would be required for colonization. The hyaluronan capsule has been demonstrated to serve a role not only in host defence evasion but also in adhesion via its interactions with CD44 molecules on the surface of human cells (Schrager et al., 1998, Schrager et al., 1996).

1.4.3.3 Fibronectin binding protein

Fibronectin is a glycoprotein found in the ECM of vertebrates and adhesins towards it, fibronectin binding proteins, are believed to be involved in the initiation of infection by some streptococcal species. These proteins are found in *S. pyogenes* (SfbI/F1) and extracellular forms have been reported in *S. equi* (FNE and SFS) which bind fibronectin (Lindmark et al., 2001, Liden et al., 2006) but are not believed to be involved in adherence and a cell bound form (FNEB) is also produced which has fibronectin binding repeats and cell wall anchoring motifs which allow for its use as an adhesin (Lannergard et al., 2005).

1.4.3.4 Hyaluronan degrading enzymes

Streptococcal cells are capable of expressing and secreting hyaluronan degrading enzymes into the surrounding environment. The production of these enzymes has been reported in streptococci Lancefield groups A, B, C and G (Allen et al., 2004) and within *Streptococcus pyogenes, Streptococcus equi* and *Streptococcus suis* there are also bacteriophage which also encode hyaluronate lyases (see section 1.4.3 streptococcal virulence factors) (Hynes and Walton, 2000, Hynes et al., 2000). Their function is for the degradation of hyaluronan of the ECM and in some cases they are also capable of the degradation of chondroitin sulphate due to similarities between the structure of these 2 compounds. Hyaluronan which forms an important part of the ECM may be degraded for the utilisation of carbon as an energy source but the function of the degrading enzymes is predicted to be more as a virulence factor as degradation allows for the dissemination of the bacteria within the host (Botzki et al., 2004, Li and Jedrzejas, 2001). A reduction in the ECM by the destruction of hyaluronan will allow for the spread of the bacteria deeper into the host tissue facilitating the establishment of a systemic infection. The enzymes are hyaluronate lyases but the mode of action is different to other non bacterial hyaluronate lyases. The degradation is by a β elimination mechanism (Rigden and Jedrzejas, 2003b) but instead of being endo-acting causing the production of a mix of size products the mode is a processive exo-acting leading to the production of unsaturated disaccharides (Pritchard et al., 1994).

It has been noted that within the Gram-positive organisms capable of producing hyaluronate lyases most, if not all, are capable of causing infection with initiation at either the mucosal membrane or on the skin surface.

1.4.3.5 Streptolysin

A major feature of GAS is the β haemolysis creating a zone of clearing around bacterial colonies grown on blood agar. The agent responsible for this haemolysis is Streptolysin S (SLS). This oxygen stable cytolysin exists in a cell bound form and targets red blood cells directly by contact with the GAS cells. SLS has a varied spectrum including neutrophils, lymphocytes, platelets as well as red blood cells. As part of its haemolytic activity SLS is capable of impairment of phagocytosis and enhancement of cytotoxicity causing increased tissue damage (Datta et al., 2005). SLS can exist as either a secreted exotoxin or as a cell bound form and both forms are suspected of being able to reduce the immune response to the bacteria by the destruction of neutrophils (Miyoshi-Akiyama et al., 2005).

A second streptolysin, Streptolysin O (SLO) has been seen in GAS and is shown to create large pores in the eukaryotic cell membranes and also acts as a vehicle for the translocation of an NADase also produced by GAS which is capable of producing cyclic ADP-ribose which acts in intracellular signalling causing many of the effects of SLO including enhanced cytotoxicity (Bricker et al., 2005).

1.4.3.6 Streptococcal pyrogenic exotoxins

Streptococcal pyrogenic exotoxins (Spes) are important virulence factors produced by streptococcal species (Anzai et al., 1999a). They are often produced in response to a low oxygen concentration suggesting that the virulence of the streptococci will increase when the bacteria enters the bloodstream from the skin or pharynx (Nakamura et al., 2004). SpeB of GAS functions as a cysteine protease and high levels of it are often associated with the clinical manifestations of STSS and ultimate mortality following infection (Tsao et al., 2006, Shelburne et al., 2005).

1.5 Bacteriophage

Recent surveys of bacteriophage have shown them to be the most populous organism of earth with approximately 10^{30} nucleic acid containing lifeforms present (Nelson et al., 2003, Chibani-Chennoufi et al., 2004). Presently there are ~5100 members of the class of bacteriophage and the International Committee for Taxonomy of Viruses (ICTV) have agreed upon a classification of 1 order, 13 families and 30 genera (Ackermann, 2003). The families can be roughly split into two groups; tailed Caudovirales comprising 96 % and the polyhedral/filamentous/pleomorphics (PFP) comprising 4 % (Ackermann, 2001). The different families are represented in Table 1.3

Shape	Family	Nucleic acid	Characteristics
Tailed	*Myoviridae*	dsDNA, linear	Contractile tail
	Siphoviridae	dsDNA, linear	Long, noncontractile tail
	Podoviridae	dsDNA, linear	Short tail
Polyhedral	*Microviridae*	ssDNA, circular	
	Corticoviridae	dsDNA, circular, superhelical	Complex capsid, lipids
	Tectiviridae	dsDNA, linear	Internal lipoprotein vesicle
	Leviviridae	ssRNA, linear	
	Cystoviridae	dsRNA, linear, segmented	Envelope, lipids
Filamentous	*Inoviridae*	ssDNA, circular	Filaments, rods
	Lipothrixiviridae	dsDNA, linear	Envelope, lipids
	Rudiviridae	dsDNA, linear	
Pleomorphic	*Plasmaviridae*	dsDNA, circular, superhelical	Envelope, lipids, no capsid
	Fuselloviridae	dsDNA, circular, superhelical	Spindle-shaped, no capsid

Table 1.3: Table showing the different families of bacteriophage and their characteristics. Modified from (Ackermann, 2003).

The tailed *Caudovirales* can be seen to be further split into 3 subfamilies: the *Myoviridae*, *Podoviridae* and *Siphoviridae* with the *Siphoviridae* comprising 61 % of tailed phages. The structure of a siphoviridae bacteriophage lambda with a long non-contractile tail is shown in Figure 1.9.

Figure 1.9: Electron micrograph of bacteriophage lambda. From http://micro.magnet.fsu.edu/cells/viruses/images/bacteriophage.jpg

During the infection process the tip of the tail fibre of the bacteriophage will interact via a protein p*J* with the LamB maltoprotein of the bacterial cell. The interaction will lead to the formation of two possible complexes; type 1, a reversible complex, and type 2, an irreversible complex. It has been reported that an irreversible complex is required for the injection of viral DNA and that the interaction between the bacteriophage and the bacterial cell will first become a type 1 complex followed by a conformational change of the phage tail leading to a type 2 complex. The bacterial receptor, LamB is a homotrimer with each monomer producing a b barrel with a water filled pore of 5-6 Å. This narrow pore is unlikely to to allow for the passage of viral DNA through it but it is believed that the binding of p*J* and LamB causes the triggering of DNA ejection from the virus into the bacterial periplasmic space. Following transit through the periplasmic space the bacteriophage DNA is then taken up by the cell through its phosphoenolpyruvate: sugar phosphotransferase system (PTS) in which transmembrane components IIC^{Man} and IID^{Man} allow for passage of the DNA through the inner membrane.

The tail fibres of different bacteriophage and associated genes demonstrate a high degree of homology not seen in other parts of the structure. The tail fibres are key for the recognition of and binding to receptor molecules on the surface of the host cell and as such need to be highly conserved. A loss of these tail fibres would render the

bacteriophage unable to bind to the host cell and as such the bacteiophage would be unable to replicate and the genome of it would be lost (Canchaya et al., 2003).

1.5.1 Prophage

Following infection the bacteriophage enters either a lytic or lysogenic life cycle. In the lytic lifecycle the bacteriophage will replicate and lead to the lysis of the host and release of the progeny bacteriophage. However in the lysogenic life cycle the bacteriophage will integrate its genetic material within the host genome and become dormant and is known as a prophage. This symbiotic relationship provides benefits for both the bacteriophage and the bacteria. The bacteriophage is supported within the host and facilitates its spread throughout a population and the bacteria will often gain genes from the bacteriophage which in many cases can cause an increase in its survival fitness. Prophage are believed to make up between 10 - 20 % of bacterial genomes (Casjens, 2003, Bossi et al., 2003, Canchaya et al., 2004) and are often responsible for the development of more pathogenic or virulent strains.

1.5.2 Bacteriophage virulence factors

The effect of bacteriophage on their bacterial host was discovered in 1927 when Frobisher and Brown observed that non-toxic streptococci could become toxic when mixed with the filtered supernatants of toxic streptococci. These supernatants are now known to have contained bacteriophage and that the bacteriophage encoded virulence factors which were transferred to the new host and converted it to a toxic form. These virulence factors which can be acquired are a wide and diverse group and can confer various different attributes to the host (Saunders et al., 2001). They can alter the adhesion properties of the cells, enhance resistance to phagocytes, encode exotoxins and affect the susceptibility to antibiotics and further infection by temperate phage (Wagner and Waldor, 2002, Tinsley et al., 2006).

1.5.2.1 Streptococcal bacteriophage encoded virulence factors

Prophage play a major role in the evolution and diversification of streptococci leading to the emergence of ever more virulent and pathogenic strains. There are two groups of prophage encoded virulence genes; the pyrogenic toxin superantigens (PTSAgs) and the non-PTSAgs. PTSAgs are secreted streptococcal pyrogenic exotoxins (Spes) and they are capable of binding simultaneously to and activating MHC Class II molecules and T cells causing the release of cytokines leading to disease symptoms such as fever, shock and dysregulation of the immune system. Non-PTSAgs are also secreted but are believed to be involved in pathogen-host interactions. These include phospholipases, hyaluronidases and DNases thought to contribute to inflammation and coagulopathy (Banks et al., 2002). DNases are also believed to have a role in improving the bacteriophage fitness. During cell lysis by the bacteriophage the viscosity will increase due to the release of DNA. This will reduce the movement of the bacteriophage to find a new host. The production of DNases will allow for the removal of the excess DNA increasing the bacteriophage movement and aiding in the location of a new host (Broudy et al., 2002).

1.6 Research objectives

Streptococcal strains have been shown to differ dramatically in their pathogenicity and this is often attributed to the role of prophage which, when in a lysogenic lifecycle, can confer additional virulence factors to the host bacterium. It has been noted that those strains of streptococci which are most virulent carry multiple prophages whilst those which are avirulent lack such prophage.

1.6.1 *S. pyogenes* Family 69 GH

Previously cloned, characterised and had its structure solved as a hyaluronate glycoside hydrolase, HylP1 from the SF370 strain of *S. pyogenes* serotype M1 is one such possible additional virulence factor. Adopting a site directed mutagenesis approach, the overall aim of the study was to investigate the role of key amino acids involved in catalysis and the production of complex crystals showing the substrate bound within the active site.

Specific objectives of this study are as follows:

1. To use site directed mutagenesis to dissect the functional significance of residues in the proposed active site.
2. To produce complex crystals using the catalytically inactive mutants to show the substrate bound within the active site.
3. To propose a mechanism of catalysis within the active site deducing the amino acids responsible.

1.6.2 *S. equi* hyaluronate glycoside hydrolases

S. equi strain 4047 has recently had its genome sequence completed by the Sanger Institute. *S. equi* is closely related to *S. pyogenes* biologically and as such is likely to have a similar prophage encoded hyaluronate glycoside hydrolase. Using bioinformatics and cloning procedures the aim of the study is to identify and clone such possible genes and biochemically characterise them. The use of an equine based pathogen may open up possible avenues for pathogenesis studies using challenge experiments which may serve as a model for *S. pyogenes* infection.

Specific aims of this study are as follows:

1. To identify, using bioinformatics, conserved regions displaying similarity to the sequence of HylP1 which may be a similar hyaluronate glycoside hydrolase.
2. To clone express and purify any identified ORF's
3. To biochemically characterise any identified proteins
4. To crystallise and obtain X-ray diffraction data for any identified proteins to compare structure with that of HylP1

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals

The chemicals, enzymes and kits used throughout this study are listed in Appendix A. Unless otherwise stated, all solutions and buffers were made up with 18.2 MΩ/cm H_2O purified using a Milli –Q Plus 185 water purification system, and stored at room temperature. For all molecular biology protocols, Sambrook and Russell (2001) was used as a general reference.

2.1.2 Equipment

A list of equipment used in this study can be found in Appendix B.

2.1.3 Media

Unless otherwise stated, all media were prepared using distilled water, and sterilised by autoclaving. Liquid media were stored at room temperature, and solid media at 4 °C. All adjustments to the pH of solutions are stated under the list of components, and were achieved using HCl or NaOH.

2.1.3.1 Liquid media

2.1.3.1.1 Luria-Bertani (LB) broth

Per L

Tryptone (Oxoid)	10 .0 g
Yeast extract (Oxoid)	5 .0 g
NaCl (Melford)	10 .0 g

pH 7.0

2.1.3.1.2 Low salt LB broth

Per L

Tryptone (Oxoid)	10 .0 g
Yeast extract (Oxoid)	5 .0 g
NaCl (Melford)	5 .0 g

pH 7.0

2.1.3.1.3 NZY Enrichment broth

Per L

NZ Amine (casein hydrolysate) (Oxoid)	10 .0 g
Yeast extract (Oxoid)	5 .0 g
NaCl (Melford)	5 .0 g
$MgSO_4.7H_2O$ (Sigma)	2 .0 g

pH 7.5

2.1.3.1.3.1 NZY^+ Enrichment broth

NZY^+ broth was used for the recovery of transformed chemically competent cells. The broth was made by the addition of 0.45 ml of supplement (stored at 4 °C) to 10 ml of NZY broth using sterile vessels and pipette tips.

NZY (supplement)

Per 45 ml

1 M $MgCl_2$ (Sigma)	12.5 ml
1 M $MgSO_4$ (Sigma)	12.5 ml
20 % (w/v) D-(+)-Glucose (Fisher)	20.0 ml

Filter sterilised.

2.1.3.1.4 SOB broth

Per 100 ml

Tryptone (Oxoid)	2.0 g
Yeast extract (Oxoid)	0.5 g
5 M NaCl (Melford)	0.2 ml
1 M KCl (Sigma)	0.25 ml

pH 7

Plus the following added aseptically prior to use

4 M $MgCl_2$ (Sigma)	0.25 ml
4 M $MgSO_4$ (Sigma)	0.25 ml

2.1.3.1.5 SOC broth

Per 100 ml SOB broth

1 M D-(+)-glucose (Fisher) (filter sterilized)	2.0 ml

2.1.3.2 Solid media

2.1.3.2.1 LB agar

For the preparation of agar plates LB media was supplemented with the following.

Per 100 ml

Agar (bacteriological agar N^o 1) (Oxoid)	2.0 g

Agar was then made soluble by autoclaving the media, and poured into Petri dishes when the media had cooled to approximately 55 °C.

2.1.3.3 Selective media

2.1.3.3.1 Antibiotics and other supplements

For antibiotic selection, appropriate volumes of stock solutions were added to liquid media that had been cooled to approximately 55 °C. The composition of the stock solutions and final concentration used in the media are given in Table 2.1. Isopropylthio-β,D-galactoside (IPTG) (Melford) was added to strains of *E. coli* carrying the *lacI*^{q} and *lacO* regions in the plasmid only, or in the plasmid and the bacterial genome to induce transcription of the recombinant gene. IPTG (100 mM; 24 mg/ml) stock dissolved in sterile 18.2 MΩ/cm H_2O) was added to liquid media to a final concentration of 1.0 mM (240 μg/ml).

Antibiotic	**Stock concentration**	**Working concentration**	**Solvent**	**Storage**
Ampicillin (Melford)	10 mg/ml	100 μg/ml	18.2 MΩ/cm H_2O	4 °C for less than 7 days or – 20 °C
Kanamycin (Melford)	10 mg/ml	50 μg/ml	18.2 MΩ/cm H_2O	– 20 °C

Table 2.1: Antibiotics used in selective media

2.1.4 Cryogenic storage of bacterial stocks

Bacterial stocks were created by mixing 0.5 ml of an overnight culture broth with 0.5 ml of sterile 50 % (v/v) glycerol. These stocks were then stored at –85 °C.

Per 100 ml

100 % (v/v) glycerol (Fisher) 50.0 ml

2.1.5 Bacterial strains and plasmids

All bacterial strains were maintained at –85 °C in 25 % (v/v) glycerol, and were used as detailed in Table 2.2. The restriction maps for the vectors used in this study are given in Appendix C.

***E. coli* strain/Vector**	**Characteristics**	**Application**	**Reference/ Supplier**
One Shot™ TOP10	F⁻ *mcrA* Δ (*mrr-hsd*RMS-*mcr*BC) ϕ80*lacZ*ΔM15 Δ*lacX74 deoR recA1 araD139* Δ(*ara-leu*)7697 *galU galK rspL endA1 nupG*	Cloning host	Invitrogen Corp.
BL21 (DE3)	F⁻*ompT hsdS*$_B$(*r*$_B$⁻*m*$_B$⁻) *gal dcm* (DE3)	Expression host	Studier *et al.*, 1986 /Novagen
pCR®-Blunt	Kan^r, Zn^r, T7 *lacZα-ccdB*	Cloning vector	Invitrogen Corp.
pET-28a	Kan^r, T7 *lac*, *lacI*^{q}	Expression vector	Novagen

Table 2.2: Bacterial strains and plasmids used in this study.

2.1.6 Molecular biology chemicals, enzymes and kits

2.1.6.1 Agarose gel electrophoresis

2.1.6.1.1 TAE Running buffer (50 x stock)

Per L

Tris Base (ultra-pure) (Melford)	242.0 g
17.51 M Glacial acetic acid (Fisher)	57.1 ml
0.5 M EDTA pH 8.0 (Duchefa)	100.0 ml

This buffer was diluted to 1 x by a 1:50 dilution prior to use in the preparation of agarose gels and use as reservoir buffer in gel tanks.

2.1.6.1.2 Bromophenol blue (6 x) sample loading buffer

Per 10 ml

Bromophenol blue (BDH Laboratory Supplies)	0.025 g
Glycerol (Fisher)	3.000 g

This buffer was diluted to 1 x with the sample prior to loading on an agarose gel.

2.1.6.1.3 Xylene cyanol (6x) sample loading buffer

Per 10 ml

Xylene cyanol (Sigma)	0.025 g
Glycerol (Fisher)	3.000 g

This buffer was diluted to 1 x with the sample prior to loading on an agarose gel.

2.1.6.1.4 Size Standards

Either Bioline Hyperladder I or NEB 1 kb ladder agarose gel size standard was diluted to a concentration of 1 μg/12 μl prior to loading on a gel (see appendix D for sizes).

2.1.6.1.5 SYBR Safe

This chemical was diluted 1:10000 in molten agarose before casting the gel.

2.1.6.2 General use chemicals

2.1.6.2.1 TE buffer

Per L

0.5 M Tris-Base (Melford) pH 7.5	20.0 ml
0.5 M EDTA (Duchefa) pH 8.0	2.0 ml

pH 7.5

2.1.6.3 Kits

The molecular biology kits employed in this study are detailed in Table 2.3. All kits and components were stored at room temperature, except the cell re-suspension buffer of plasmid purification kits, which was maintained at 4 °C.

Kit	Application
Qiagen Miniprep	Purification of recombinant plasmids
Qiagen Maxi prep	Large scale preparation of cloning vectors
Qiagen Qiaquick gel extraction kit	Purification of DNA from agarose gels
Qiagen DNeasy	Purification of genomic DNA

Table 2.3: Molecular biology kits and applications.

For a list of components of each kit, refer to Appendix A.

2.1.6.4 Enzymes

All enzyme buffers and reaction co-constituents were maintained at – 20 °C, in 50 % (v/v) glycerol. A list of general enzymes used in this study is shown in Table 2.4. However, a full list of all enzymes and constituents of reaction buffers is in Appendix A.

Enzyme (stock concentrations)	**Co-constituents**
KOD DNA polymerase (Novagen)	Amplification buffer (1 x): Patented
Taq DNA polymerase (5 U/µl) (New England Biolabs)	Reaction buffer (1 x): 10 mM Tris-HCl pH 9.0, 1.5 mM $MgCl_2$, 50mM KCl, 0.1 % Triton® X-100
Restriction enzymes: *Bam* HI, *Nde* I (New England Biolabs)	A list of buffer constituents used for each restriction enzyme is given in Appendix A3
RNase (10 mg/ml) (Qiagen)	Added to Qiagen Resuspension Buffer: 10 mM Tris-HCl pH 7.5, 1 mM EDTA, to a final concentration of 100 µg/ml RNase
T_4 DNA ligase (133 U/µl) (New England Biolabs)	Ligation buffer (1 x): 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 µg/ml BSA
Dpn I Restriction enzyme (10 U/µl) (New England Biolabs)	Used unbuffered

Table 2.4: Enzymes and co-constituents.

2.1.7 Primer Sequences

2.1.7.1 Amplification of SEQ2045

The primer sequences given in Table 2.5 were used for the amplification of SEQ2045 (available from http://www.sanger.ac.uk/Projects/S_equi/S_equi.art and the triple stranded β helix of HylP1 (GenBank Accession number AAK33657).

2.1.7.2 Creation of HylP1 mutants

The primer sequences given in Table 2.7 were used for the creation of site-directed mutants of HylP1 and SEQ2045. The primers were designed so that the bases on either side of the desired mutated bases were fully complementary to both the sense and antisense strands of wild type HylP1 and SEQ2045.

Primer name	Primer sequence	Primer length (mer)
SEQ2045f	5'CATATgTCAAAAgAAgTTgCATC3'	23
SEQ2045rs	5'ggATCCTATTTTTTTAgTATgAgTTTTTTTAAC3'	33
TSβHK124f	5'CATATGAAATTAAATCTCAAAggTggCgTTATg3'	33
TSβHT339rs	5'ggATCCTTACCAgTAAgggTTCCAgTC3'	27

Table 2.5: Primer sequences used for the amplification of SEQ2045 and the triple stranded β helix of HylP1. The genomic DNA used as the template was *Streptococcus equi 4047 for SEQ2045 and Streptococcus pyogenes* SF370 (ATCC 700294) for the triple stranded β helix.

Primer pair	PCR product amplified	Primary (1°) annealing temperature (°C)	Secondary (2°) annealing temperature (°C)
SEQ2045f + SEQ2045rs	SEQ2045 gene (stop)	40	50
TSβHk124f + TSβHT339rs	TSβH gene (stop)	53	58

Table 2.6: Primer pairs used for amplification of genes and the primary and secondary annealing temperatures of the pairs

Primer name	Sequence	Length (mer)	Tm	% GC	% mismatch
HylP1I136Af	5'gggTggAgCggTCAATgCggACTTgTCgTCTACCAg3'	36	79.8	61	8
HylP1I136Ar	5'CTggTAgACgACAAgTCCgCATTgACCgCTCCACCC3'	36	79.8	61	8
HylP1D137Af	5'gTggAgCggTCAATATTgCgTTgTCgTCTACCAgAC3'	36	78.5	53	6
HylP1D137Ar	5'CTCTggTAgACgACAACgCAATATTgACCgCTCCAC3'	36	78.5	53	6
HylP1Y149Af	5'GgTgCTggTgTTgTTgTCgCgTCTgACAATgATACCAgTg3'	40	78.4	53	8
HylP1Y149Ar	5'CACTggTATCATTgTCAgACgCgACAACAACACCAgCACC3'	40	78.4	53	8
HylP1Y149Ff	5'GTgCTggTgTTgTTgTCTTTTCTgACAATgATACCAgTg3'	39	79.2	44	3
HylP1Y149Fr	5'CACTggTATCATTgTCAgAAAAgACAACAACACCAgCAC3'	39	79.2	44	3
SEQ2045D171Af	5'ggAggAgCAATTAACATTgCgATgTCTAAATCggAA3'	36	74.0	42	6
SEQ2045D171Ar	5'CCTTCCgATTTAgACATCgCAATgTTAATCgCTCCT3'	36	75.2	45	6
SEQ2045Y183Af	5'ggTgCTggTgTTgTCgCgTCTAACAATgATACC3'	33	73.4	52	9
SEQ2045Y183Ar	5'CCATCACTggTATCATTgTTAgACgCgACAACAACA3'	36	73.2	45	8
SEQ2045Y183Ff	5'gTgCTggTgTTgTTgTCTTTTCTAACAATgATACCA3'	36	75.7	39	3
SEQ2045Y183Fr	5'CATCACTggTATCATTgTTAgAAAAgACAACAACAg3'	36	74.5	36	3

Table 2.7: Primer sequences used for the generation of site directed mutants

2.1.8 Vectors

For use in the cloning of PCR products, pCR®-Blunt was used, and for the subsequent expression of cloned open reading frames pET-28a was used. Restriction maps of the multiple cloning regions of the stated vectors are provided in Appendix C.

2.1.9 Solutions, resins and buffers for protein analysis and purification

2.1.9.1 SDS – PAGE chemicals

2.1.9.1.1 12 % (w/v) Acrylamide resolving gel components

40 % (w/v) solution (37.5:1 acrylamide: bisacrylamide) (Fisher)	3.0 ml
Solution B*	2.5 ml
18.2 MΩ/cm H_2O	4.5 ml
10 % (w/v) Ammonium persulphate (Fisher)	50.0 µl
Tetramethylethylenediamine (TEMED) (Fisher)	10.0 µl

*Solution B

Per 100 ml	
2 M Tris Base (Melford) pH 8.8	75.0 ml
10 % (w/v) Sodium dodecyl sulphate (SDS) (Melford)	4.0 ml

2.1.9.1.2 4% (w/v) Acrylamide stacking gel components

40 % (w/v) solution (37.5:1 acrylamide: bisacrylamide) (Fisher)	0.5 ml
Solution C*	1.0 ml
18.2 MΩ/cm H_2O	2.5 ml
10 % (w/v) Ammonium persulphate (Fisher)	30.0 µl
TEMED (Fisher)	10.0 µl

*Solution C

Per 100 ml	
1 M Tris Base (Melford) pH 6.8	50.0 ml
10 % (w/v) SDS (Melford)	4.0 ml

2.1.9.1.3 SDS - PAGE running buffer 10 x (Stock)

Running buffer was made up at 10 x stock concentration and diluted 1:10 before use.

Per L

Tris Base (Melford)	30.3 g
Glycine (Melford)	144.0 g
SDS (Melford)	10.0 g

The pH of the Tris Base and glycine was adjusted to 8.8 in a volume of ~900 ml prior to the addition of SDS.

2.1.9.1.4 SDS - PAGE sample buffer

Per 10 ml

60 mM Tris Base (Melford) pH 6.8	0.6 ml
50 % (v/v) Glycerol (Fisher)	5.0 ml
10 % (w/v) SDS (Melford)	2.0 ml
14.4 mM β-Mercaptoethanol (Sigma)	0.5 ml
1 % (w/v) Bromophenol blue (BDH)	1.0 ml

Stored at –20 °C in 0.5 ml aliquots.

2.1.9.1.5 Solublising SDS - PAGE sample buffer

Per 10 ml

SDS - PAGE sample buffer	7.6 ml
Urea (Melford)	2.4 g

Stored at 4 °C.

2.1.9.1.6 Protein size standards

High molecular weight range (M.W. 36, 45, 55, 66, 84, 97, 116 and 205 kDa)
Low molecular weight range (M.W. 20, 24, 29, 36, 45 and 66 kDa)

Lyophilised standards were reconstituted with 100 µl of 18.2 MΩ H_2O to give a final concentration of ~2.0-3.5 mg/ml, which was aliquoted into 4 µl amounts and stored at –20 °C. Refer to Appendix D for a list of proteins used to produce the appropriate sized bands.

2.1.9.1.7 Coomassie blue gel stain solution

Per L	
Coomassie Blue R-250 (Fisher)	1.0 g
Glacial Acetic acid (Fisher)	100.0 ml
Methanol (Fisher)	450.0 ml

Coomassie blue gel stain solution is recycled for re-use by filtering through filter paper to remove any small pieces of gel.

2.1.9.1.8 Coomassie gel destain solution

Per L	
Glacial Acetic acid (Fisher)	100.0 ml
Methanol (Fisher)	100.0 ml

Coomassie gel destain solution is recycled for re-use by filtering through activated charcoal on a filter paper funnel to remove the Coomassie blue stain.

2.1.9.1.9 Western blot transfer buffer

Per L	
Tris base (Melford)	5.82 g
Glycine (Melford)	2.93 g
Methanol (Fisher)	200 ml

2.1.9.1.10 Phosphate buffered saline (PBS)

Per L

1 M K_2HPO_4 (Fisher)	8.0 ml
1 M KH_2PO_4 (Fisher)	1.98 ml
KCl (Sigma)	0.2 g
NaCl (Melford)	8.0 g

2.1.9.1.11 PBS-Tween solution

Per L

PBS	999.5 ml
Polyoxyethylenesorbitan monooleate (Tween 80) (Sigma)	0.5 g

2.1.9.1.12 Marvel-PBS-Tween solution

Per 100 ml

Marvel powdered milk (Premier International Foods)	5.0 g
PBS-Tween solution	95.0 ml

2.1.9.1.13 BCIP/NBT

Concentrated stocks (10x) of both the buffer solution and the reagent were diluted to 1x in 18.2 MΩ H_2O.

2.1.9.2 Protein purification

2.1.9.2.1 Cell resuspension buffer (post-induction)

Per L

HEPES	4.77 g
NaCl	29.22 g
Imidazole	0.68 g

pH 7.4

2.1.9.2.2 Purification resins

2.1.9.2.2.1 Ni^{2+} resin

Fast flow chelating sepharose (GE Healthcare)

2.1.9.2.2.2 Gel filtration resin:

Superdex 200 prep grade (GE Healthcare)

All resins were stored in 20 % (v/v) ethanol.

2.1.9.2.3 Ni^{2+} affinity purification buffers (Automated purification, gradient flow)

2.1.9.2.3.1 Start solution:

Per L	
HEPES (Melford)	4.77 g
NaCl (Melford)	29.22 g
Imidazole (Acros)	0.68 g
pH 7.4	

2.1.9.2.3.2 Elution solution:

Per L	
HEPES (Melford)	4.77 g
Imidazole (Acros)	68.0 g
pH 7.4	

2.1.9.2.4 Gel filtration buffers

2.1.9.2.4.1 For HylP1

Per L	
HEPES (Melford)	4.77 g
$CaCl_2$ (Sigma)	29.4 g
pH 7.4	

2.1.9.2.4.2 For SEQ2045

Per L	
HEPES (Melford)	4.77 g
NaCl (Melford)	11.9 g
pH 7.4	

2.1.9.2.5 Buffer exchange buffers (Post-gel filtration)

This was achieved using 10 kDa concentrators (Vivaspin).

2.1.9.2.5.1 For HylP1

Per L

HEPES (Melford)	2.38 g
$CaCl_2$ (Sigma)	7.35 g

pH 7.4

2.1.9.2.5.2 For SEQ2045

Per L

HEPES (Melford)	0.595 g

pH 7.4

2.1.9.3 Reagents for Bradfords assay

BSA (Sigma) was made at a concentration of 10 mg/ml and stored at –20 °C, for subsequent dilution in assay. Bradford Reagent (Brilliant blue G in phosphoric acid and methanol) (Sigma) was stored at 4 °C.

2.2 Methods

2.2.1 Substrate preparation

2.2.1.1 Dialysis

Dialysis of substrates was achieved in cellulose membrane tubing (size 25 mm x 16 mm; retains > 90 % Cytochrome C (M.W. 12 400 Da); Sigma) and secured with dialysis tubing closures (gripping length 50 mm; Sigma).

Sodium hyaluronate (Sigma) and potassium hyaluronate (Sigma) were dialysed against 18.2 MΩ H_2O/cm (3 x 4 L).

2.2.1.1.1 Preparation of dialysis tubing

The dialysis tubing was cut into pieces of convenient length and boiled for 10 min in a large volume of 18.2 MΩ/cm H_2O. From this point onward, the tubing was always handled with gloves. The tubing was then rinsed thoroughly inside and out with 18.2 MΩ/cm H_2O and once cool, was then stored at 4 °C in 18.2 MΩ/cm H_2O.

2.2.1.1.2 Preparation of dialysed potassium/sodium hyaluronate

Potassium or sodium hyaluronate (Sigma) was prepared in the presence of the metal ion chelator EDTA (Duchefa). Potassium or sodium hyaluronate (1 g) was added to 18.2 MΩ/cm H_2O (90 ml) and autoclaved. EDTA at 500 mM pH 8 was added to a final concentration of 50 mM and this solution was then stirred for 5 h at room temperature. Precipitates were then removed by centrifugation for 15 min at 14 000 x g and supernatants were subsequently dialysed three times in 4 L 18.2 MΩ/cm H_2O overnight. The dialysed substrate was then aliquotted in 20 ml volumes into Petri dishes, frozen at –20 °C and lyophilised overnight and stored at room temperature for future use. Prior to kinetic analysis, the substrate was resuspended in 18.2 MΩ/cm H_2O.

2.2.2 Microbiology methods

2.2.2.1 Sterilisation

Unless stated otherwise, all solutions and apparatus were sterilised by autoclaving at 121 °C, with a pressure of 1.05 bar for 20 min. Solutions that could not be autoclaved were filter sterilised through a 0.2 μm Ministart® filter unit (Sartorius), attached to a suitable sterile syringe (Plastipak®, Becton Dickinson).

2.2.2.2 Growth of bacteria harbouring plasmids

2.2.2.2.1 Growth of bacteria for standard/crude plasmid purification

For high copy number plasmids, *E. coli* strain TOP10 clones were grown in 25 ml universals containing 5 ml of LB media. For low copy number plasmids, *E. coli* strain TOP10 clones were grown in 250 ml conical flasks containing 50 ml of LB media, supplemented with the appropriate antibiotic. The cultures were inoculated from an agar plate, glycerol stock or liquid culture using a sterile wire loop, and grown overnight at 37 °C with orbital shaking at 200 rpm.

2.2.2.2.2 Growth of bacteria for starter cultures

E. coli strains TOP10 or BL21 (DE3) were grown in 25 ml sterile glass universals containing 5 ml of LB media supplemented with the appropriate antibiotic. The cultures were inoculated from an agar plate, glycerol stock or liquid culture using a sterile wire loop, and grown overnight at 37 °C with orbital shaking at 200 rpm.

2.2.2.2.3 Growth of bacteria for the expression and subsequent purification of protein

A starter culture of *E. coli* BL21 (DE3) was grown (see section 2.2.2.2.2) and was asceptically added at a 1 % (v/v) ratio to a conical flask containing LB media supplemented with the same antibiotic used in the starter culture. The volume of LB media used depended on the amount of protein required to perform appropriate biochemical or structural techniques. Table 2.8 shows the two scales of expression used.

Scale of expression	Volume of LB media	Size of conical flask	Volume of starter culture added
Small	50 ml	250 ml	0.5 ml
Large	1000 ml	2500 ml (Baffled)	10.0 ml

Table 2.8: Volumes and vessels used in the growth of expression strains.

After addition of the starter, the culture was grown at 37 °C with orbital shaking at 200 rpm until the culture reached an OD_{600} of ~0.6. The OD of the culture was checked by transferring 1 ml of culture to a cuvette and reading the OD in a spectrophotometer previously zeroed with LB media and set to 600 nm. Once at 0.6, the cells were induced by the addition of 100 mM IPTG to a final concentration of 1 mM IPTG. Post induction, the culture was incubated at 30 °C overnight with orbital shaking at 100 rpm.

2.2.2.3 Plating bacteria

Prior to plating, LB agar plates were surface dried by placing open face down at 65 °C for 10 min. A glass spreader was sterilised by immersion in 70 % (v/v) ethanol, followed by removal of excess ethanol by passing quickly through a blue Bunsen flame, and used to spread 50-200 µl of bacterial suspension evenly over the surface of the agar. When all the liquid had been absorbed into the agar, the plate was incubated inverted at 37 °C for ~ 16 h.

2.2.2.4 Preparation of chemically competent cells

For all strains of *E. coli*, chemically competent cells were freshly prepared using the procedure detailed here and was essentially by the method of (Cohen et al., 1972). Firstly, an agar plate of the glycerol stock of the required cell type was prepared and incubated overnight at 37 °C. After overnight growth, 50 ml of prewarmed LB media in a 500 ml conical flask was inoculated using a wire loop. The culture was then allowed to grow at 37 °C with shaking at 200 rpm up to an OD_{600} of 0.35~0.4. At this point the culture was placed on ice for 30 min. After incubation on ice, the culture was pelleted by centrifugation at 4000 x g for 10 min at 4 °C and the supernatant discarded. The cell pellet was then re-suspended (by gentle vortex never more than 3 sec at a time off ice) in 15 ml of ice-cold 80 mM $MgCl_2$, 20 mM $CaCl_2$ solution, and the cells were then re-centrifuged at 4000 x g for 10 min at 4 °C and the supernatant

removed. The cell pellet was re-suspended (by gentle vortex never more than 3 sec at a time off ice) in 1 ml of ice- cold 100 mM $CaCl_2$. The cells were then retained on ice for 1h and were then used immediately or snap frozen (in liquid N_2) and stored at -85°C subsequent to the addition of 50 % (v/v) glycerol to a final concentration of 15 % (v/v).

2.2.2.5 Preparation of electrocompetent cells

Firstly an agar plate of the glycerol stock of the required cell type was prepared and incubated overnight at 37 °C. After overnight growth, 50 ml of prewarmed low salt LB media in a 500 ml conical flask was inoculated using a wire loop. The culture was then grown at 37 °C with shaking at 200 rpm until the cells reached an optical density of 0.5~0.7 at 600 nm. The culture was then placed on ice for 30 minutes. The culture was then centrifuged at 4000 x g for 15 min at 4 °C and the supernatant discarded. The pellet was then gently resuspended (by gentle vortex never more than 3 sec at a time off ice) in 25 ml of ice cold filter sterilized 18.2 MΩ/cm H_2O. The suspended cells were then centrifuged as before and the supernatant discarded. The cell pellet was then resuspended (by gentle vortex never more than 3 sec at a time off ice) in 12 ml of ice cold filter sterilized 18.2 MΩ/cm H_2O. The centrifugation and resuspension was repeated twice more but with the resuspension volume decreasing each time with the next volume being 5 ml and this was followed by resuspension in 2 ml. A final centrifugation and resuspension step followed with a volume of 0.2 ml of ice cold filter sterilized 18.2 MΩ/cm H_2O and the cells then aliquoted into pre-chilled microcentrifuge tubes in 40 µl amounts ready for transformation.

2.2.3 Bioinformatic analysis

2.2.3.1 Preliminary bioinformatics

Protein sequences were analysed by bioinformatics firstly for their molecular weight and pI using the EXPASY (expert protein analysis system) identification and characterisation tool (www.expasy.org). The protein was also studied for the presence of signal peptide cleavage sites using SignalP 3.0 server available at www.cbs.dtu.dk/services/SignalP.

2.2.3.2 Alignment of sequences

All alignments were performed using ClustalW software from www.ebi.ac.uk/Tools/clustalw. Multiple sequences were entered and the alignment ran. Following this, conserved and semi conserved residues could be determined.

2.2.3.3 BLAST analysis

To identify possible similar sequences to genes or proteins, BLAST searches were performed. DNA blasts (BLASTN) were used when looking at gene sequences and protein blasts (BLASTP) were used when looking at protein sequences. Both were used from the National Centre for Biotechnology Information (www.ncbi.nlm.nih.gov/BLAST).

2.2.3.3.1 *S.equi* and *S.pyogenes* BLAST searches

For prophage analysis, BLAST searches were performed on single specific organisms and strains. For this specific BLAST tools were used. When searching *S. equi* the BLAST was performed at www.sanger.ac.uk/Projects/S_Equi and when searching *S. pyogenes* the BLAST was performed at http://www.ncbi.nlm.nih.gov//genomes/geblast.cgi?bact=off&gi=178

2.2.4 DNA Methods

2.2.4.1 General DNA methods

2.2.4.1.1 Primer design

Primers were designed (table 2.5) according to the following rules

Must be more than 18 nucleotides in length

Must end in either a G or a C

The pairs must have annealing temperatures within 5 °C of each other

Annealing temperatures were calculated by the following equation:

69.3 + (0.41 x % G/C) – (650 / primer length)

2.2.4.1.2 Polymerase chain reaction (PCR) protocol

For the amplification of SEQ2045

Primers were designed to anneal to the 5' and 3' ends of the sequence encoding the protein (see table 2.5). The primers were also designed to contain restriction endonuclease recognition sites both upstream and downstream of the target sequence. These engineered digestion sites allowed for the ORF to be inserted in the correct frame into a pET expression vector. A PCR product of the gene was amplified from genomic DNA of *Streptococcus equi* 4047. The PCR product generated was cloned into PCR-Blunt and used for subsequent cloning into pET-28a vector so that the encoded products contained an *N*- terminal His_6 tag.

The reactions for PCR were set up in sterile microcentrifuge tubes (0.2 ml) as detailed in Table 2.9.

Tube component	**Volume added**	**Final concentation in PCR reaction.**
Sterile 18.2 MΩ/cm H_2O	34 µl	
Forward primer (12.5 µM)	1 µl	0.25 µM
Reverse primer (12.5 µM)	1 µl	0.25 µM
KOD reaction buffer 10 x	5 µl	1 x
S. equi 4047 genomic DNA 0.1 µg/µl	1 µl	0.002 µg / µl
$MgSO_4$ (25 mM)	2 µl	1 mM
dNTPs (2 mM) (each)	5 µl	0.2 mM (each)
KOD polymerase (3 U/µl)	1 µl	0.06 U / µl

Table 2.9: Reaction components for PCR.

The tubes were capped and centrifuged briefly to bring the contents to the bottom of the tube. The Eppendorf MasterCycler™ machine was then programmed to proceed as shown in Table 2.10. A gradient PCR protocol was used where the first five cycles

of PCR were carried out at the primary annealing temperature and the subsequent 25 cycles of PCR were carried out at the secondary annealing temperature. The primary annealing temperature was calculated to be 5 °C lower than the lowest T_m of the primer pair (excluding the primer sequence that did not anneal to the genomic DNA). The secondary annealing temperature was calculated to be 5 °C lower than the lowest T_m of the primer pair (all of the primer sequence). The primary and secondary annealing temperatures are given in table 2.6.

Temperature (°C)	Duration (min:sec)	Cycles
94.0	02:00	First denaturation
94.0	00:15	5
1° annealing temperature	00:30	
68.0	02:00	
94.0	00:15	25
2° annealing temperature	00:30	
68.0	02:00	
68.0	10:00	Final extension
10.0	Hold	

Table 2.10: Reaction conditions for PCR.

After the PCR reactions had cooled to 10 °C, a 10 µl aliquot was removed and analysed by agarose gel electrophoresis (Section 2.2.3.1.4) to check for successful amplification.

2.2.4.1.3 Site-directed mutagenesis (SDM) protocol

For the production of *HylP1* and *SEQ2045* site directed mutants, a pET-28a clone containing *HylP1* or *SEQ2045* was used as the template. Primers were designed and are given in Table 2.7. The reaction components for the amplification of the mutants is shown in Table 2.11 and the order the components appear in the table is the same as the order of addition to the PCR tube.

Tube component	Volume added	Final concentration in PCR reaction.
Sterile 18.2 MΩ/cm H_2O	40.5 µl	
Forward primer (125 ng/µl)	1 µl	2.5 ng/µl
Reverse primer (125 ng/µl)	1 µl	2.5 ng/µl
KOD reaction buffer 10 x	5 µl	1 x
Template DNA (500 ng/µl)	1 µl	10 ng/µl
dNTPs (25 mM) (each)	0.5 µl	0.25 mM (each)
KOD polymerase (3 U/µl)	1 µl	0.06 U/µl

Table 2.11: Reaction components for PCR of site-directed mutants. For the recipe of *KOD* reaction buffer (10 x), refer to Appendix.

The Eppendorf MasterCycler™ machine was then programmed to proceed as shown in Table 2.12

Temperature (°C)	Duration (min)	Cycles
95.0	00:30	1
95.0 55.0 68.0	00:30 01:00 13:12	16
10.0	Hold	

Table 2.12: Reaction conditions for PCR of site-directed mutants.

2.2.4.1.4 Restriction digests

All restriction endonuclease (New England Biolabs) digests were performed according to the manufacturers instructions. Unless otherwise stated, each µg of DNA was cut with 5 Units of enzyme at 37 °C for > 1.5 h. Heat inactivation was carried out at 65 °C for 10 min when required. BSA was added to the reaction mix at a final concentration of 0.1 mg/ml when required, and the appropriate reaction buffer was also added to a final concentration of 1 x. Care was taken when preparing digests that the glycerol concentration did not exceed 10 % (v/v) in the final reaction, in case enzyme specificity was affected or star activity was caused.

Digests with two or more restriction enzymes were carried out simultaneously if the enzymes had high enough (> 50 %) activity in the same buffer. A list of restriction endonucleases used in this study and corresponding reaction buffers are given in Appendix A. The digested samples were then subjected to agarose gel electrophoresis. The required DNA fragment was recovered from the gel using a Qiagen Qiaquick gel extraction kit and used for ligation.

2.2.4.1.5 Agarose gel electrophoresis

Agarose gels were prepared by heating agarose in 1 x TAE buffer until the solid had completely dissolved. The amount of agarose added depended on the size of the gel to be cast. All gels in this study were 1 % (w/v) agarose in a total volume of 30 ml, except those gel cast for purification purposes, which were made in a volume of 50 ml in order to give the wells a greater capacity.

Once dissolved, the solution was allowed to cool to ~ 60 °C, the appropriate volume of SYBR Safe was added, and the solution poured into a mini gel casting tray avoiding the introduction of air bubbles. The appropriate sized comb was then inserted into the solution (for analytical purposes, a 12-toothed comb able to hold 15 µl in each well was used, and for preparative purposes, a 5-toothed comb able to hold 50 µl in each well was used) and the gel allowed to set for > 20 min.

After the gel had set, the gel tray was placed horizontally in an agarose gel electrophoresis tank, submerged in 1 x TAE buffer. The comb was then carefully removed and 10 µl samples were prepared by the addition of 2 µl of sample buffer (Bromophenol blue 6 x or Xylene cyanol 6 x), and added to the wells. A 6 µl volume of size standard was also added to an unoccupied well in order to determine the size of electrophorised DNA fragments. The samples were then electrophoresed at 100 mA (200 V) for approximately 40 min to ensure good separation of DNA fragments.

2.2.4.1.6 Visualisation of DNA and photography of agarose gels

After electrophoresis, the gel was visualised using a gel documentation system (Bio-Rad Gel Doc 2000 using Quantity One™ software). Hard copies of the gel picture were produced using a Mitsubishi Video Copy Processor (Model P91 attached to the gel doc system), with Mitsubishi thermal paper (K65HM-$_{CE}$ / High density type, 110 mm x 21 m).

2.2.4.1.7 Preparation of plasmid DNA (pDNA)

2.2.4.1.7.1 Standard preparation of pDNA

Purification of plasmid DNA from transformed *E. coli* strain TOP10 was achieved using a Qiagen Miniprep plasmid purification kit. The method by which the purification was performed is detailed here.

A 3 ml volume of bacterial cell culture was pelleted in a 1.5 ml micro-centrifuge tube by repeated centrifugation at 14000 x g for 60 sec, and the supernatant discarded. The pellet was then pulsed and residual supernatant removed. The pellet was then re-suspended by vortexing in 250 µl of buffer **P1** (cell re-suspension buffer). Buffer **P2** (250 µl) (cell lysis buffer) was then added and the sample was mixed by gentle inversion 5-6 times. Buffer **N3** (300 µl) (neutralization buffer: for the precipitation of proteins and chromosomal DNA) was then added to the sample, and the sample was again mixed by gentle inversion 5-6 times. Precipitated material was then removed by centrifugation at 14000 x g for 10 min, and the supernatant was applied to a Qiagen filter column placed into a 2 ml collecting tube. The filter column was then centrifuged at 14000 x g for 1 min and the flowthrough discarded. Buffer **PB** (500 µl) was then applied and centrifuged (14000 xg for 1 min) through the column and the flowthrough discarded. The column was then washed with 750 µl of buffer **PE** with ethanol (14000 x g for 1 min) and the flowthrough discarded. The column was then centrifuged dry to remove all traces of ethanol (14000 x g for 1 min). The Qiagen filter column was transferred into a sterile 1.5ml micro-centrifuge tube with the lid removed to allow for centrifugation and the DNA was eluted by centrifugation at 14000 x g for 1 min after the addition of 50 µl of buffer **EB** (elution buffer). The components of each buffer are given in Appendix A.

2.2.4.1.7.2 Preparation of plasmid DNA for sequencing

For the preparation of plasmid DNA for sequence analysis, the standard preparation of plasmid DNA method was employed, but the plasmid DNA was eluted in 2 x 50 µl volumes of sterilized 18.2 MΩ/cm H_2O. The concentration of the DNA was then calculated and a volume equivalent to 1 µg was dried using a centrifugal evaporator ready for sequencing. Following centrifugal evaporation the DNA was sent to MWG Biotech for sequencing.

2.2.4.1.8 Agarose gel purification

Purification of DNA fragments from agarose gels was achieved using a Qiagen Qiaquick gel extraction kit. The method by which the purification was performed is detailed here.

Prior to removal of the fragment, the gel was 'briefly' visualised under low intensity UV and etching using a clean scalpel blade indicated the location of the fragment to be excised. The section of gel containing the fragment was then excised using a scalpel, and inserted directly into a pre-weighed 1.5 ml micro-centrifuge tube. The micro-centrifuge tube was then re-weighed and the weight of the gel slice it contained was calculated. For every 100 mg of agarose gel, 300 µl of buffer **QG** was added to the tube. The gel was then dissolved by incubation at 50 °C for 10 min with a brief vortex every 2-3 min. Once dissolved, 100 ml of isopropanol for every 100 mg of gel was added. The sample was then loaded into a filter column and placed into a 2 ml collecting tube. The sample was then centrifuged at 10000 x g for 60 sec, and the flowthrough discarded. Residual traces of agarose were removed by the addition of 500 µl buffer **QG**, which was centrifuged at 14000 x g for 60 sec and the flowthrough discarded. The filter column was then washed by the addition of 750 µl of buffer **PE** and centrifuged at 14000 x g for 60 seconds and the flowthrough discarded. This was followed by a second centrifugation at 14000 x g for 60 sec of the empty filter column and collecting tube, in order to remove any residual ethanol from the filter. The filter column was then transferred into a sterile 1.5 ml micro-centrifuge tube with the lid removed and 30 µl of **EB** buffer was added precisely to the top of the filter membrane. The sample was then centrifuged at 14000 x g for 60 sec. The filter was then discarded and the DNA that was eluted into the micro-centrifuge tube was retained for subsequent use. To check the purification had

worked, 2 µl was removed from the eluted sample, and made up to 10 µl with 18.2 MΩ/cm H_2O. This sample was then analysed by agarose gel electrophoresis. The components of each buffer are given in Appendix A.

2.2.4.1.9 Preparation of vector DNA

A glycerol stock of *E. coli* TOP10 known to harbour the vector was streaked out onto an appropriate antibiotic selective media and grown over night at 37 °C. A single colony was then used to inoculate a starter culture of between 2 and 5 ml of appropriate antibiotic selective broth. This was then grown with shaking overnight and the following day diluted 1:1000 into an appropriate volume of antibiotic selective media. For low copy number plasmids this was 500 ml and for high copy number plasmids it was 100 ml. The culture was then grown overnight at 37 oC with shaking. The cells were then harvested by centrifugation at 6000 x g for 15 min at 4 °C. The supernatant was decanted and the pellet resuspended fully in 10 ml of buffer **P1** (resuspension buffer). The cells were lysed by the addition of 10 ml buffer **P2** (cell lysis buffer) and then neutralized by the addition of 10 ml buffer **P3** (neutralization buffer). The sample was then centrifuged at ≥ 20000 x g for 30 min at 4 °C. The supernatant containing the plasmid DNA was centrifuged agin at the previous settings but for 15 min. During this time a QIAGEN-tip 500 was equilibrated with 10 ml buffer **QBT** and allowed to flow by gravity. The centrifuged supernatant was applied to the equilibrated column and then the column was washed with 2 x 30 ml of buffer **QC**. The DNA was eluted from the column by the addition of 15 ml of buffer **QF**. The resulting DNA was precipitated by adding 10.5 ml of room temperature isopropanol and decanted into 1.5 ml microcentrifuge tubes which were then centrifuged at ≥15000 x g for 30 min at 4 °C. The supernatants were discarded and each pellet washed with 300 µl of 70 % ethanol and centrifuged at ≥15000 x g for 10 min at 4 °C. The supernatant was carefully decanted without disturbing the DNA pellet and each pellet allowed to air dry until all traces of ethanol were gone. Each pellet was redissolved in 10 µl of TE buffer (see section 2.1.6.2.1) and was dissolved pooled and quantified.

2.2.4.1.10 Spectrophotometric quantification of DNA

DNA concentration was quantified accurately by scanning through 200-300 nm using a UV spectrophotometer.

DNA was quantitatively determined by preparing a dilution of DNA in a 70 µl total volume. The dilution was then transferred to a 50 µl quartz cuvette and scanned in a UV vis spectrophotometer (previously zeroed with 18.2 MΩ/cm H_2O) between the wavelengths of 200 and 300 nm. Optimal wavelength for the absorbance of DNA was assumed to be 260 nm, and 280 nm was assumed to be the optimal wavelength for contaminating protein, and as double stranded DNA at 50 µg/ml has an A_{260} reading of 1.0 the concentration and purity of the DNA was determined as detailed below.

[dsDNA] µg/µl = (A_{260}) x 50 x dilution / 1000.
DNA purity ratio = (A_{260})/ (A_{280})

A DNA purity ratio of 1.5-1.8 was accepted as minimal protein contamination.

2.2.4.1.11 Ligation of DNA

For ligations into pCR®-Blunt a 10:1 molar insert to vector ratio was required and for ligation into pET vector a 2:1 molar insert to vactor ratio was required. Concentrations were determined by running samples on an agarose gel and using the analysis tool of the Bio-Rad Gel Doc 2000 Quantity One™ software.

2.2.4.1.11.1 Ligation into pCR®-Blunt vector

Ligation of blunt-ended PCR products into linearised pCR®-Blunt was performed by the addition of the components listed in Table 2.13. A 10:1 molar insert-to-vector ratio was used.

Component	Volume
PCR®-Blunt vector (linearised; 25 ng)	1 µl
10 x Ligation buffer (with ATP)	1 µl
Blunt PCR product	1-5 µl
18.2 MΩ/cm H_2O	To total 10 µl
T_4 DNA Ligase (4U/µl)	1 µl

Table 2.13: Reaction components for ligation into pCR®-Blunt vector.

This reaction was then incubated at 16 °C overnight prior to transformation into TOP10 chemically competent cells.

2.2.4.1.11.2Ligation into pET vector

Ligation into pET vector (Appendix C) was performed by the addition of the components listed in Table 2.14.

Component	Volume
PET vector cut with appropriate restriction enzymes (50 – 100 ng/µl; Lab stock)	2 µl
10 x Ligation buffer	1 µl
Fragment cut from pCR®-Blunt vector and gel purified (50 ng/µl)	2 µl
18.2 MΩ/cm H_2O	4 µl
T_4 DNA Ligase (3 U/µl)	1 µl

Table 2.14: Reaction components for ligation into pET vector.

After addition of reaction components, the ligation was incubated at 4 °C overnight prior to transformation into TOP10 cells.

2.2.4.1.12 Transformation of competent cells

2.2.4.1.12.1 Chemical transformation using *E. coli* TOP10 and BL21 (DE3) cells

Transformations were performed using prepared chemically competent cells in 50 µl aliquots. Firstly, the DNA to be transformed was added and mixed gently with a pipette tip and incubated with the cells on ice for 20 min. The volume of DNA added from a ligation was 2 µl, and from a plasmid preparation was 1 µl. The cells were then transferred to a water bath at 42 °C and heat shocked for 90 sec, prior to a second incubation on ice for 2 min. The cells were then recovered by the aseptic addition of 200 µl of room temperature NZY^+ enrichment broth, and incubation at 37 °C for 1 hour shaking (horizontally for maximum aeration) at 200 rpm.

Recovered cells were then aseptically plated on media appropriate to the vector transformed. Bacteria transformed with pET plasmids were plated on the appropriate antibiotic. Plates were then incubated inverted for ~ 16 h at 37 °C to allow the growth of bacterial colonies.

2.2.4.1.12.2 Electroporation transformation using *E. coli* TOP10 cells

Transformations were performed using prepared electrocompetent cells. 40 µl aliquots of cells were placed on ice and 2 µl of ligation were added and mixed gently with a pipette tip. The cells and DNA mix were incubated on ice for at least one minute. The mix was then transferred to a ice cold sterile 2 mm electroporation cuvette and the contents tapped gently down to the bottom of the cuvette until it was entirely covering the bottom of the cuvette. The outside of the cuvette was dried thoroughly and the cuvette placed into a BioRad ShockPod and the lid closed. The BioRad PulseXcell was set to 2.5 kV and the voltage applied across the cuvette. Immediately following electroporation 1 ml of SOC media (see section 2.1.3.1.5) was added to the cuvette and transferred using a sterile glass pipette to a sterile 1.5 ml microcentrifuge tube. The cells were incubated at 37 °C with shaking at 200 rpm for at least 1 hour and then plated onto an appropriate antibiotic LB agar plate. Plates were then incubated inverted for ~16 hours at 37 °C.

2.2.4.2 Construction of mutant plasmids of *HylP1* and *SEQ2045* by site directed mutagenisis (SDM)

Mutagenic oligonucleotide primers were designed with a length of 25-45 bases, a GC content ≥ 40 °C and a T_m (melting temperature) of ≥ 78 °C with the mutation designed centrally in the primer. The T_m for each primer was determined as follows:

$$T_m = 81.5 + 0.41\ (\%\ GC) - 675 / N - \%\ mismatch$$

Key:
- % GC = Percentage of guanine and cytosine residues in the primer
- N = Primer length in bases
- % mismatch = Percentage of the primer non complementary at the point of mutation

Mutant plasmid was then generated by PCR. The PCR template DNA (i.e. the methylated, non-mutated supercoiled DNA) was then digested at 37 °C for 1 h by adding 1 µl of *Dpn* I (10 U/µl) restriction enzyme directly to the cooled PCR reaction and mixed gently. The *Dpn* I treated DNA was then transformed into *E. coli* TOP10, and plasmid DNA made from one of the transformants was sent to MWG Biotech for sequence analysis in order to confirm the mutation.

2.2.5 Protein methods

2.2.5.1 SDS - PAGE

SDS – PAGE was conducted according to the method of Laemmli (Laemmli, 1970) using equipment from the BioRad Mini Protean III kit

SDS - PAGE gels were cast between two clean glass plates of dimensions (10.1 x 7.2 and 10.1 x 8.2 cm) with the larger plate having 0.75 mm spacer ridges attached to them. The plates were clamped together and checked to ensure the bottoms were flush, prior to securing the gel with vertical downward pressure in a casting clamp stand on a rubber gasket. The components for a resolving gel were then mixed together in a plastic universal, and then pipetted into the space between the plates, until the solution was approximately 2 cm below the top of the smallest plate. In

order to prevent inhibition of polymerisation by oxygen and to produce a straight top edge, a small volume of water was then added to the top of the solution. The resolving gel was then allowed to polymerise for ~ 20 min. Once set, the water was removed, the plates dried using blotting paper (care was taken not to touch the surface of the resolving gel), and the components of the stacking gel were mixed, and pipetted up to the top of small plate. A 10-toothed comb (1.1 x 0.75 cm) was then immediately inserted into the gap between the plates. The stacking gel was then allowed to polymerise for ~ 20 min prior to removal of the comb to reveal the wells, which were immediately rinsed with water to remove any gel debris and unpolymerised acrylamide. The gels were then placed into the electrophoresis module with the small plates facing inwards to create a central reservoir then placed vertically into the tank. SDS-PAGE (1 x) running buffer was then carefully poured into the tank ensuring no bubbles were created and that the level of buffer in the space between the two gels was lower than the level on the outside of the gels but was filling the wells and covering the top of the gel.

Samples were then prepared by the addition of 5 µl of SDS-PAGE loading buffer to 20 µl of sample, which was then pulsed briefly prior to boiling for 3 min. Size standards were prepared in the same way, except 8 µl of loading dye was added to 4 µl of size standard. All samples and standards were then centrifuged at 14000 x g for 1 min prior to loading 20 µl of all samples and 10 µl of all standards into the wells using a Hamilton syringe. The pair of gels were then electrophorised at 120 milliamps, 200 volts for ~ 50 mins, or until the bromophenol blue dye had migrated off the bottom of the gel.

2.2.5.1.1 Visualisation of protein bands and photography of SDS-PAGE gels

After electrophoresis, the gels were removed from the apparatus and stained for 10 min in Coomassie blue stain, followed by immersion in destain until crisp band images were visible against a clear background. The gels were then photographed using a gel documentation system. Hard copies of the gel picture were produced using a Mitsubishi Video Copy Processor, with Mitsubishi thermal paper (Appendix B for equipment).

2.2.5.2 Isolation of cell free extract (CFE)

To obtain a CFE, firstly induced cells were concentrated by centrifugation for 10 min at 5500 x g, after which the supernatant was decanted and the cells were re-suspended in 5-20 ml of cell re-suspension buffer. Cells were then lysed in 5 ml volumes by sonication on ice at amplitude 14 for 1 min in 10 sec bursts with 10 sec intervals. The lysate was then decanted into centrifuge tubes and centrifuged at 4 °C for 20 min at 24000 x g. After centrifugation the soluble fraction was decanted into a plastic universal and maintained on ice as CFE.

2.2.5.3 Protein purification

Recombinant proteins used in this study were purified by a number of different methods. In all cases, the purity of the protein was assessed by SDS-PAGE. Immobilised metal affinity chromatography (IMAC) was used to purify recombinant proteins, under native conditions, which contained a stretch of 6 Histidine residues at the N-terminus (His_6 tag). The metal affinity matrix used was Sepharose™ chelating fast flow resin.

2.2.5.3.1 Affinity purification using a fast flow Ni^{2+} column

Recombinant proteins fused to a His_6 tag were purified under native conditions. For the purification of protein using the IMAC technique, the column was equilibrated by passing though 100 ml of buffer.

For the high yield purification of protein using the automated FPLC® system (Appendix B for automotive parts) gradient elution technique, the column was prepared by adding ~ 50 ml Sepharose™ chelating fast flow resin to a C series column (Pharmacia Biotech). The column adapters were securely attached and the column was connected to the FPLC® system. Once connected, 1 M $NiSO_4$ solution was passed through the resin until all the resin was a uniform colour. The column was then washed by passing through ~ 200 ml of Start buffer (see section 2.1.9.2.3.1) using the FPLC® machine. Once the column had been equilibrated, CFE (10 –20 ml) was then loaded onto the column. The column was then washed with 100 ml of Solution A at a flow rate of 5 ml/min generated by the FPLC gradient pump. Finally, the protein was eluted over 50 min at a flow rate of 5 ml/min using a linear gradient of imidazole (Elution buffer) (see section 2.1.9.2.3.2) extending from 10 mM to 500 mM, using Solution A as a diluent. The presence of the target protein was

determined by monitoring at A_{280} using an inline UV spectrometer connected to the FPLC, and samples were collected in 5 ml volumes on a fraction collector. To confirm purity a 20 µl aliquot from each 5 ml fraction thought to contain the target protein analysed by SDS - PAGE and the pure fractions pooled.

2.2.5.3.2 Gel filtration chromatography

Purified protein from an affinity chromatography column was exchanged by concentration into gel filtration buffer. The gel filtration column, of dimensions 1.6 cm (i.d.) x 60 cm (L), with a bed volume of 120 ml was prepared by equilibration (using the FPLC® system) at 1 ml/min with 120 ml of gel filtration buffer (see section 2.1.9.2.4). Protein sample, 0.5 ml, was then loaded onto the column by injection and eluted over 120 min with gel filtration buffer running at 1 ml/min. The elution point of the target protein was identified by A_{280} using the inline UV spectrophotometer, and samples were collected on a fraction collector in 5 ml volumes. To confirm purity, a 20 µl aliquot from each 5 ml fraction thought to contain the target protein was analysed on a SDS-PAGE gel and the pure fractions were pooled.

2.2.5.3.3 Concentration and buffer exchange of protein

Concentration of protein was performed at 4 °C by centrifugation at 4000 x g in 10 kDa Vivaspin 6ml cut-off concentrator units (Viva Science). Protein was also buffer exchanged by dilution within concentrator units.

2.2.5.4 Determination of protein concentration

2.2.5.4.1 Bradford method

The amount of protein present in samples was determined by the method of (Bradford, 1976).

Protein determination by the Bradford method involved the construction of 10 dilutions of BSA in 500 µl aliquots ranging from 1-10 µg/ml. Bradford's solution (500 µl) was then added to each dilution, and mixed. Each standard was then measured in a glass cuvette at A_{595} against a blank containing 500 µl of 18.2 MΩ/cm H_2O mixed with 500 µl of Bradford's solution. The glass cuvette was washed with methanol after every measurement, and all measurements were completed within 40

min of initially adding the Bradford's solution. A standard curve was then constructed. Dilutions of the protein to be determined were then made in 500 µl volumes and measured after the addition of Bradford's solution in the same way as the standards. Those values that fell on the standard curve were then used to calculate protein concentration, using the equation below.

Protein conc mg/ml = (Conc obtained from the curve) x (dilution of sample) / 1000

2.2.6 Assay methods

2.2.6.1 Assay methods for HylP1 and SEQ2045

2.2.6.1.1 General considerations for kinetic analysis

- All kinetics assays was performed using the same set of calibrated Gilson pipettes.
- The same aliquot of purified enzyme was used in kinetic experiments.
- All enzyme assays were performed in triplicate and the mean values were presented on the graphs with error bars to represent the range for each data set.

2.2.6.1.2 232 nm lyase assay

A reaction volume of 495 µl was made containing potassium hyaluronate and 50 mM buffer. The reaction was transferred to a 500 µl quartz cuvette and placed into the UV vis spectrophotometer. The UV vis spectrophotometer was under remote control by an attached PC using Vision 32 software. The spectrophotometer was also attached to a 37 ^{o}C circulating water bath to allow the cuvette to be warmed. After a 5 minute incubation period to warm the reaction in the cuvette the absorbance at 232 nm was measured in real time and plotted by the software. Once a baseline was established 5 µl of appropriately diluted enzyme (0.2 µg/µl) was added and mixed briefly by pipetting. The absorbance at 232 nm was monitored by the software and plotted to create a real time change in absorbance graph. Once complete the rate of the reaction could be calculated using the line of best fit tool in the software using the tangent to curve of the graph.

2.2.6.1.3 Lyase inhibitor assay

The effect of Vitamin C as an inhibitor of HylP1 and SEQ2045 was determined using a dinitrosalicylic acid (DNSA) assay. A reaction volume was set up containing 2 mg/ml potassium hyaluronan, 50 mM of an appropriate buffer and varying concentration of Vitamin C. The reaction was warmed at 37 °C for 5 min. Following incubation, an appropriate volume of enzyme was added (equivalent to 1 μg) and the reaction mixed by vortexing. To monitor the reaction 150 μl aliquots removed at various time points (0-64 min) and added to 150 μl DNSA reagent (1 % (w/v) DNSA, 0.2 % (v/v) phenol, 1 % (w/v) NaOH, 0.002 % (w/v) glucose, 0.05 % (w/v) $NaSO_3$) to terminate the reaction. For time point 0 min, 150 μl of the reaction mix was removed prior to the addition of the enzyme. The tubes were then boiled for 20 min, placed on ice for 10 min and finally equilibrated to room temperature before reading the absorbencies. The absorbencies were determined at 570 nm in a 96-well plastic microtitre plate (200 μl of each reaction was measured) using a plate reader (Appendix B). Reducing sugar release was determined from a standard curve. Standard curves were produced each time an assay was carried out. The standard curves were generated by the addition of 0-1.0 mg/ml of glucose (final concentration; 10 mg/ml stock concentration) to 2 mg/ml (final concentration) of soluble potassium hyaluronate in 50 mM of an appropriate buffer to a final volume of 150 μl. DNSA reagent (150 μl) was added to the standard curve and then processed in the same manner as the enzymes assays (post-incubation). Assays were carried out at 37 °C, unless otherwise stated.

2.2.6.1.4 Digests for HPAEC analysis

Enzyme (HylP1 or SEQ2045), diluted to an appropriate concentration, was added to 2 mg/ml potassium hyaluronate (final concentration) and 50 mM ammonium actetate buffer, pH 6.5, mixed by vortexing and 150 μl aliquots removed at 0, 10, 20, 40, and 60 min. To terminate the reaction, each digest was boiled for 5 min before being analysed by HPAEC. For 0 time, 150 μl of the reaction mix was removed prior to the addition of the enzyme.

2.2.6.1.5 Determination of K_M, k_{cat} and specific activity

For the determination of kinetic parameters of HylP1 and SEQ2045 against potassium hyaluronan, varying substrate concentrations between 4 and 0.25 mg/ml (final concentrations) were chosen and assays performed as detailed in section 2.2.6.1.2. In order to ascertain if these data reflected true Michaelis- Menten kinetics, a Lineweaver-Burk plot was plotted and if a straight line was observed the data were accepted as true.

2.2.6.1.5.1 Calculation of K_m and K_{cat}

Calculation of Kcat required the use of extinction coefficient for the double bond formed at the non-reducing end of the polysaccharide; this value was taken as 5500 M^{-1} cm^{-1}. Km was calculated by reciprocalising the X intercept in the Lineweaver Burk plot to give Km in mg ml^{-1}. Kcat is calculated through the attainment of a value for Vmax. Kcat is calculated by first using the Beer-Lambert law:

$C = \Delta A/E$

C = moles of product produced per min per L

ΔA = change in absorbance per 1 min

E = Extinction coefficient of the double bond (5500 M^{-1} cm^{-1})

This then gives the value for C. C is expressed as moles of product released per min per L but the reaction volume is only 500 ml so C is divided by 2000 to obtain the moles of product released per min per 500 ml. This value is then converted to mmoles of product by multiplying by 106 and this value is then defined as U. U is then divided by the total amount of enzyme added to the reaction in mg. This is the value for Vmax which is; units of enzyme activity per min per mg. A further conversion is needed to convert the mmoles of product released per min per mg of enzyme to the mmoles of product realeased per min per mmole of enzyme by multiplying by the molecular weight of the enzyme in mg. This is then converted to a value of mmoles per min per mmole to mmoles per sec per mmole (Kcat) by dividing by 60.

2.2.6.1.6 Determination of pH optimum

For the determination of pH optimum of HylP1 and SEQ2045 a reaction containing the substrate potassium hyaluronan (2 mg/ml final concentration), enzyme and 50 mM of the following buffer was tested: sodium acetate pH 4.0, 4.5, 5.0, 5.5; ammonium acetate pH 6.0, 6.5, 7.0, 7.5; glycine pH 8.0, 8.5, 9.0.

2.2.6.1.7 Determination of temperature optimum and thermostability

Temperature optimum for HylP1 and SEQ2045 were determined under standard assay conditions, except the reaction was carried out at a range of temperatures (27, 37, 47, 57, 67 and 77 °C).

The thermostability of HylP1 and SEQ2045 was assessed by pre-incubating the enzyme at 27, 37, 47, 57, 67, and 77 °C for 20 min using a thermocycler prior to assaying the samples under standard conditions.

2.2.6.1.8 Determination of divalent ion requirement

The effects of divalent cations on activity was assessed using the standard reaction conditions, except divalent ions $CaCl_2$, $BaCl_2$, $CoCl_2$, $CuCl_2$, $MgCl_2$, $MnCl_2$, $NiCl_2$, $SrCl_2$, $ZnCl_2$ were all including in the reaction at 1 mM.

2.2.7 Western blot analysis

To perform Western blot analysis of SEQ2045, firstly a 20 µl aliquot of purified protein was run on a 12 % (w/v) polyacrylamide gel. Following electrophoresis the gel was soaked in 20 ml of transfer buffer for 10 min. At the same time 1 piece of nitrocellulose membrane with the same dimensions as the gel and 4 pieces of blotting paper were also soaked in 20 ml and 100 ml of transfer buffer respectively. The base plate of a BioRad semi dry western blotter was dampened with transfer buffer and the blot assembled as follows: 2 layers of soaked blotting paper, nitrocellulose membrane, gel, 2 layers of blotting paper ensuring no air bubbles were caught beneath the layers as this would cause blank patches in the blot. The upper plate of the blotter was then dampened with transfer buffer and the blotter assembled taking care not to reverse the polarities. The blotter was run at 60 mA per gel for 1 h.
Following blotting, the membrane was stained in Ponceau S protein stain to locate the bands in the gel and to ensure complete transferral of proteins. The stain was then

washed from the membrane using PBS-Tween and the membrane blocked overnight in ~ 20 mL PBS-Tween containing 5 % (w/v) marvel milk powder.

The following day the primary antibody (in the form of horse sera) was diluted 1 in 1000 in 20 mL PBS-Tween containing 5 % marvel milk powder and left with gentle agitation for 2-3 hours. The membrane was then washed in ~ 20 mL PBS-Tween for 5 min and this wash was repeated 5 times to ensure complete removal of the primary antibody. The secondary alkaline phosphatase conjugated anti horse antibody was diluted 1 in 30000 in 20 mL PBS-Tween containing 5 % marvel milk powder and incubated with gentle agitation for 2-3 hours followed by the same washing procedure as used for the primary antibody. The membrane was than developed using a BCIP/NBT kit (Zymed) whereby 1 ml of buffer solution was diluted in 8 ml of 18.2 MΩ/cm H_2O followed by the addition of 1 ml of reagent and visualised and electronically stored using Quantity One software.

2.2.8 Mode of action by HPAEC analysis

HPAEC was performed by the method described by Lauder (Lauder et al., 2000). The samples were placed in a Dionex 0.5 ml sample vial with 25 mM $NaBH_4$. The samples were run through a CarboPac PA1 column 250 mm x 4 mm using a Dionex system (Appendix B). Eluent A: 0.1 M NaOH, eluent B: 1.3 M NaCl in 0.1 M NaOH were run at 1 ml/min as follows. A 12 min isocratic phase of 98 % A, 2 % B; a 50 min linear gradient increasing the concentration of B to 46 %; an 8 min linear gradient increasing the concentration of B to 87 %; a 6 min linear gradient increasing the concentration of B to 100 %; a 4 min isocratic phase of 100 % B. Following each run a wash step of 15 min of 0.1 M NaAcetate in 0.1 M NaOH was included to prepare the column for the next run. Programs for running and data collected were all managed using Dionex PeakNet software.

2.2.9 Circular Dichroism of HylP1 mutants

Recombinant protein was purified to homogeneity by IMAC and gel filtration chromatography. It was then concentrated and washed into 10 mM HEPES pH 7 and 50 mM $CaCl_2$ to a final concentration of 20 mg/ml. This was then incubated at a final concentration of 4 mg/ml with various concentrations of guanidinium hydrochloride ranging from 0-5 M for 30 minutes at room temperature. CD spectra (260-210 nm) averaged from 15 scans were collected at 20°C with a bandwidth setting of 3nm, a scanning speed of 100 nm/min and a 1 mm pathlength cuvette using a J-810 spectropolarimeter (Jasco U.K Ltd)

3 Results

3.1 HylP1

3.1.1 Contextualisation of results

HylP1 was previously cloned and biochemically and structurally characterised as a glycoside hydrolase of hyaluronan (Smith, 2004). The elucidated structure of the enzyme showed a trimerised protein which was intricately woven in the central region to give a triple stranded β helix (Figure 3.1). Within the triple stranded β helix there are three long apparent grooves which are flanked on each side by a different single member of the trimerised protein.

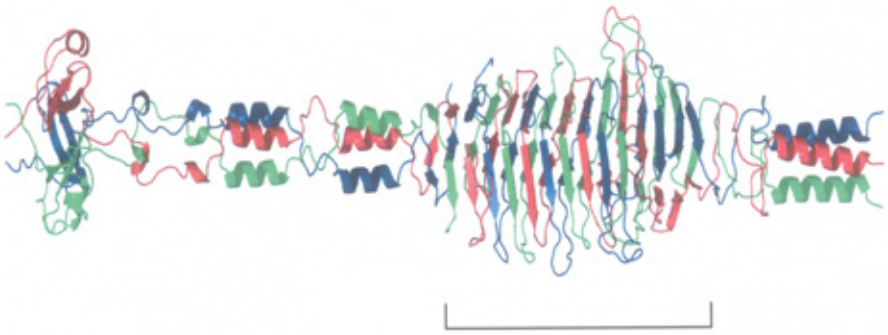


Figure 3.1: 3D structure of HylP1. Different coloured strands represent the three monomers of the trimeric structure. The triple stranded β helix region is indicated

This triple stranded β helix structure element has been previously observed in other proteins all of which make up the tail fibres of bacteriophage. The T4 short tail fibre protein gp12 (Thomassen et al., 2003) and cell-puncturing device gp5 (Kanamaru et al., 2002) and the core structure of the *Escherichia coli* K1 bacteriophage K1F endo, 2-8, sialidase (Stummeyer et al., 2005). The gp5 and the endo, 2-8, sialidase are known to be involved in phage infection of the host bacterium by degradation of the peptidoglycan and the polysialic capsule, respectively.

Previously characterised as a hyaluronidase (glycoside hydrolase of hyaluronate) by Smith (2004), as it was originally assigned to glycoside hydrolase CAZy family GH69, literature searching and comparisons with other bacteriophage tail fibre proteins capable of degradation of polysaccharides indicated the possibility that HylP1 was incorrectly characterised and was in fact a hyaluronate lyase.

HylP1 was therefore expressed and purified (figure 3.2) and preliminary assays were conducted according to a standard lyase method detailed in section 2.2.6.1.2. Activity was seen confirming that HylP1 was a hyaluronate lyase and not a hyaluronidase. A hyaluronate lyase can be assayed for reducing sugar production and increase in absorbance at 232 nm as it produces both a reducing residue and an unsaturated residue at the newly fomed non-reducing end of the polysaccharide. Contrary to this a true hyaluronidase can only be assayed by the measurement of the production of reducing sugars as no unsaturated residue is produced. This discovery created the need for full recharacterisation as a hyaluronate lyase before any other studies could be conducted.

This chapter will describe the biochemical characterisation of HylP1 of *S. pyogenes*. Data will also be shown regarding site directed mutagenesis of HylP1 allowing for the elucidation of the amino acid residues contributing to catalysis. Following from this, a characterisation of a truncated form of the protein, comprised of the suspected catalytic region, is also shown.

The results show that HylP1 is catalytically active on hyaluronate and behaves as a polysaccharide lyase, cleaving the substrate by a β elimination mechanism. Biochemically, HylP1 has an optimum pH of 6.5 and demonstrated highest activity at 37 °C. There was no absolute requirement for any of the divalent ions tested. HylP1 was determined to have a Km of 1.47 mg ml^{-1} and a kcat of 7.2 s^{-1} against dialysed potassium hyaluronate. HPAEC analysis showed the enzyme to have an

endo-cleaving mechanism with the smallest product of digestion being a tetramer. Site directed mutagenesis of suspected catalytic residues indicated a role in catalysis for D137, I136 and Y149 but subsequent circular dichroism analysis showed the loss of activity observed following mutation of I136 was due to incorrect or incomplete folding of the protein resulting in a conformational change.

A truncated form of the protein, containing only the region which forms the triple stranded β helix seen in the trimerised protein was cloned and expressed. This was found to behave in a similar manner to the full length protein with a Km of 11.1 mg ml^{-1} and a kcat of 0.53 s^{-1}. Site directed mutagenesis of the same residues mutated in the full length protein again yielded a protein with significantly reduced activity.

3.1.2 Protein expression and purification

3.1.2.1 Expression and purification of HylP1 recombinant protein

To allow for the biochemical characterisation of HylP1 as a hyaluronate lyase the *E. coli* BL21 DE3 cells harbouring the plasmid encoding the native recombinant protein (Smith, 2004) was grown at 37 ^{o}C until mid exponential phase. Expression of HylP1 was induced by the addition of 1 mM IPTG (final concentration) and the culture incubated at 30 ^{o}C overnight (section 2.2.2.2.3.) The following day cells were harvested by centrifugation and a cell free extract prepared as described in section 2.2.4.2. The recombinant protein was purified from the cell free extract by immobilised metal affinity chromatography (IMAC) using a chelating sepharose column and an automated gradient elution technique (section 2.2.4.3.1). The protein was checked for purity by SDS-PAGE as shown in figure 3.2 (section 2.2.4.1) and the concentration determined by Bradford's assay (section 2.2.4.4.1) before being subjected to biochemical analysis (section 2.2.5)

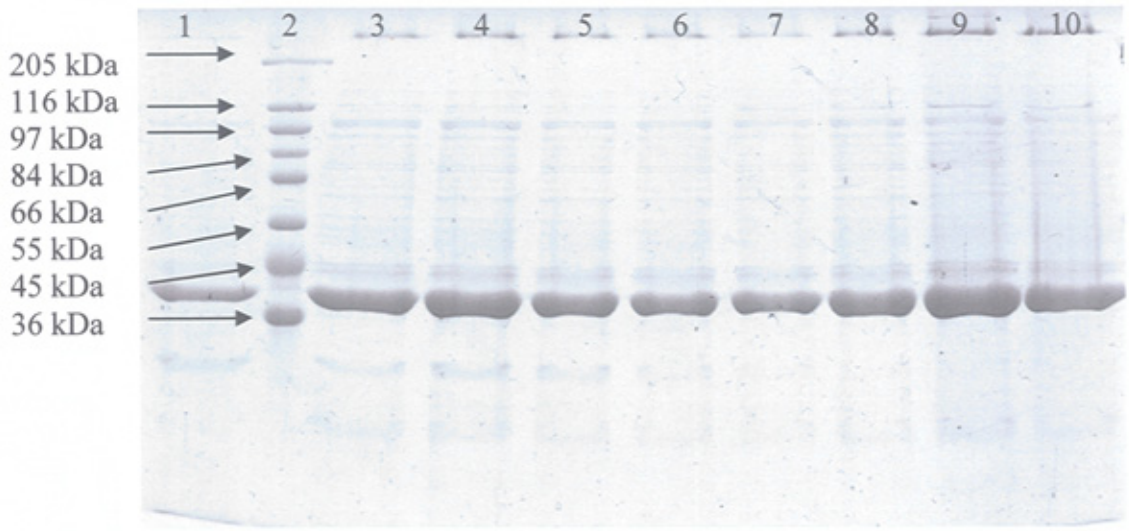


Figure 3.2: 12 % SDS PAGE showing IMAC purified fractions of HylP1.

Lanes 1, 3-10; individual fractions of purified HylP1, lane 2; high molecular weight marker (205, 116, 97, 84, 66, 55, 45, 36 kDa)

3.1.3 Kinetic analysis of HylP1

3.1.3.1 Michaelis-Menten parameters for HylP1

For the investigation of the activity of HylP1 against hyaluronate as a lyase, the enzyme was assayed for an increase in absorbance at 232 nm against dialysed potassium hyaluronate. For absolute kinetic analysis, 4 to 5 substrate concentrations (determined empirically) were chosen and reactions performed in triplicate. To ascertain if the data reflected true Michaelis-Menten kinetics, a Lineweaver-Burk plot from the three replicates was constructed as seen in figure 3.3 and used for the determination of a K_m of 1.47 mg ml^{-1} and a K_{cat} of 7.2 s^{-1}.

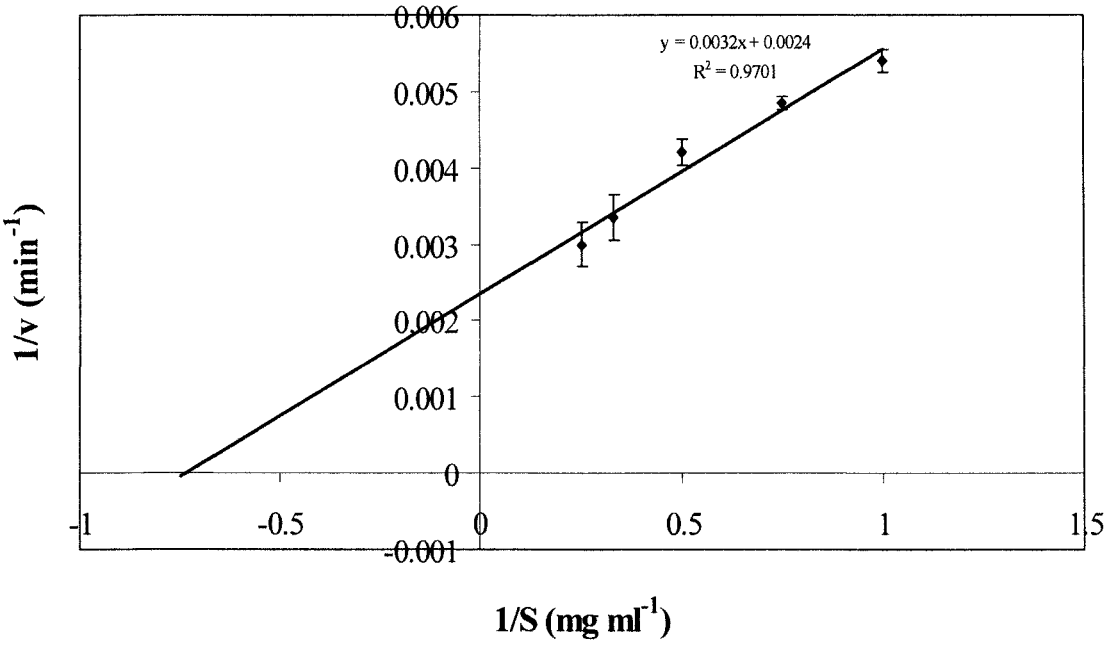


Figure 3.3: Lineweaver-Burke plot for HylP1 against dialysed potassium hyaluronate. The reactions were performed in triplicate as detailed in section 2.2.6.1.2. Each reaction had 1 µg of enzyme added and was performed by the method of Baker *et al*, (2002). A linear trendline was added and error bars represent the standard deviation from the mean.

3.1.4 Biochemical and biophysical parameters of HylP1

3.1.4.1 Determination of pH optimum

HylP1 was shown to have an optimum pH of 6.5 against dialysed potassium hyaluronate. A range of buffers were used to determine the pH optimum. Assays were carried out at 37 °C and the rate measured spectrophotometrically at 232 nm. Figure 3.4 shows the effect of pH on enzyme activity with the curve being constructed from 3 independent data sets.

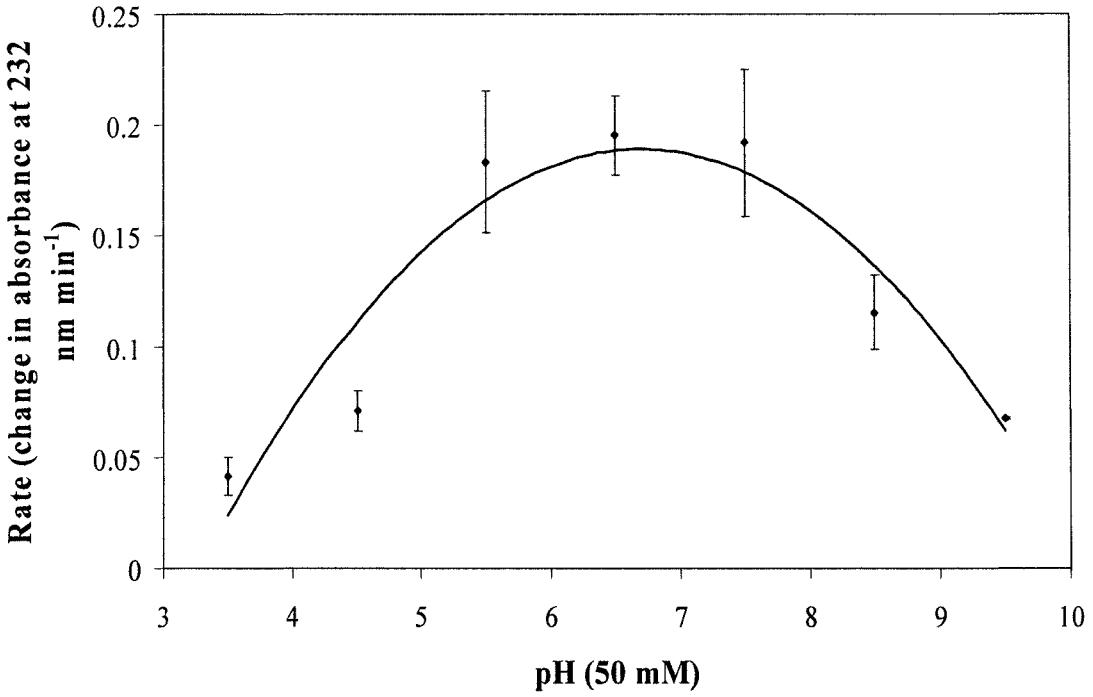


Figure 3.4: The effect of pH on the rate of enzyme activity against dialysed potassium hyaluronate. The reactions were performed in triplicate as detailed in section 2.2.6.1.2. Each reaction had a final substrate concentration of 2 mg ml^{-1} and 1 μg of enzyme added and was performed by the method of Baker *et al*, (2002). A polynomial trendline was added and error bars represent the standard deviation from the mean.

3.1.4.2 Determination of temperature optimum and thermostability

At 37 °C, HylP1 demonstrated the highest amount of activity against dialysed potassium hyaluronate when added chilled to a pre-warmed reaction mixture as seen in figure 3.5. At temperatures above and below 37 °C the enzyme exhibited reduced activity.

Upon pre-incubation for 20 min at various temperatures prior to assay, HylP1 retained highest activity at 27 °C and activity was significantly reduced at temperatures above 47 °C as seen in figure 3.6.

The effect of temperature on activity and stability was determined from 3 independent data sets.

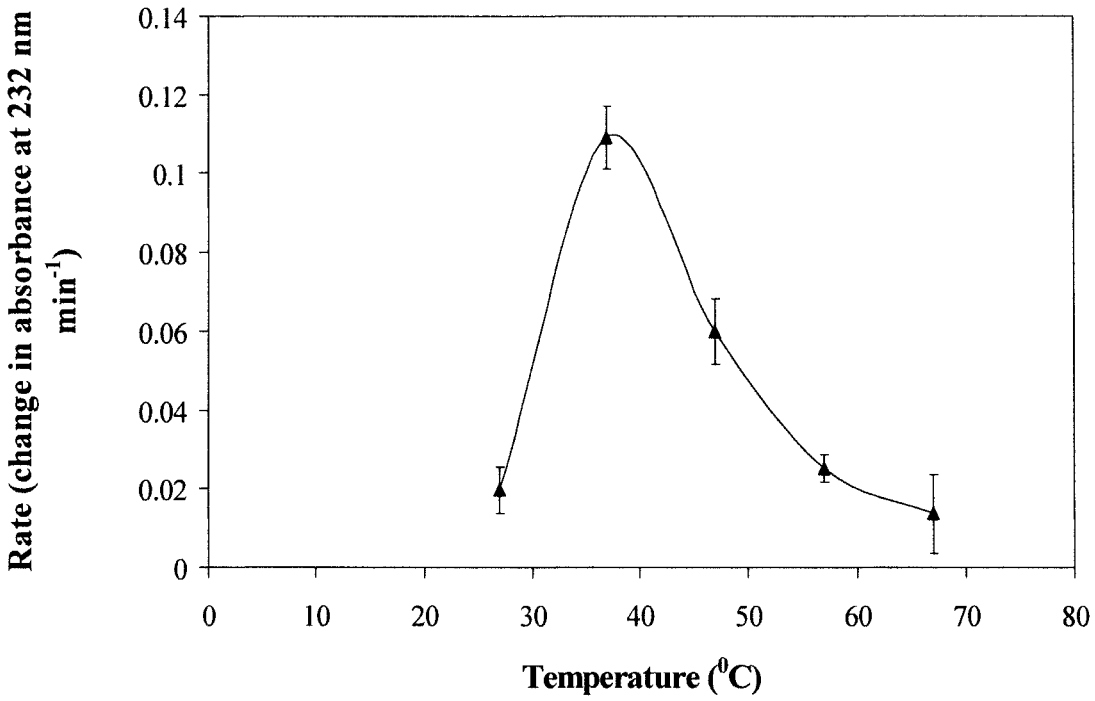


Figure 3.5: The effect of temperature on the rate of enzyme activity against dialysed potassium hyaluronate. The reactions were performed in triplicate as detailed in section 2.2.6.1.2. Each reaction had a final substrate concentration of 2 mg ml^{-1} and 1 μg of enzyme added and was performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

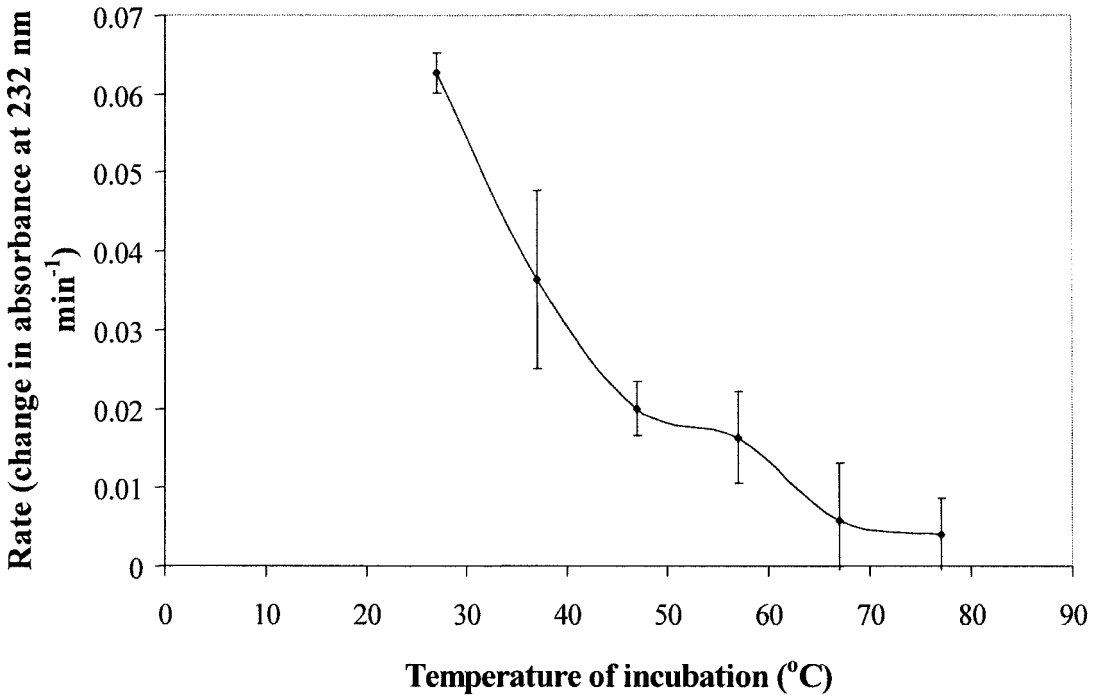


Figure 3.6: The effect of temperature on enzyme stability against dialysed potassium hyaluronate. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} and 1 μg of enzyme added after pre incubation of the enzyme for 20 min at 27, 37, 47, 57, 67 and 77 °C and were performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

3.1.4.3 Determination of divalent ion requirement

When HylP1 was tested with a range of divalent cations, it was observed that the enzyme had absolutely no requirement for any of the ions tested. The data for the effect of divalent ions on activity is presented in figure 3.7. Activity was seen to increase slightly when $SrCl_2$ and $BaCl_2$ were added and a decrease was observed when $CuCl_2$ was added.

A range of increased concentrations of $CaCl_2$ were also added and HylP1 again showed no requirement for $CaCl_2$ at higher concentrations as seen in figure 3.8

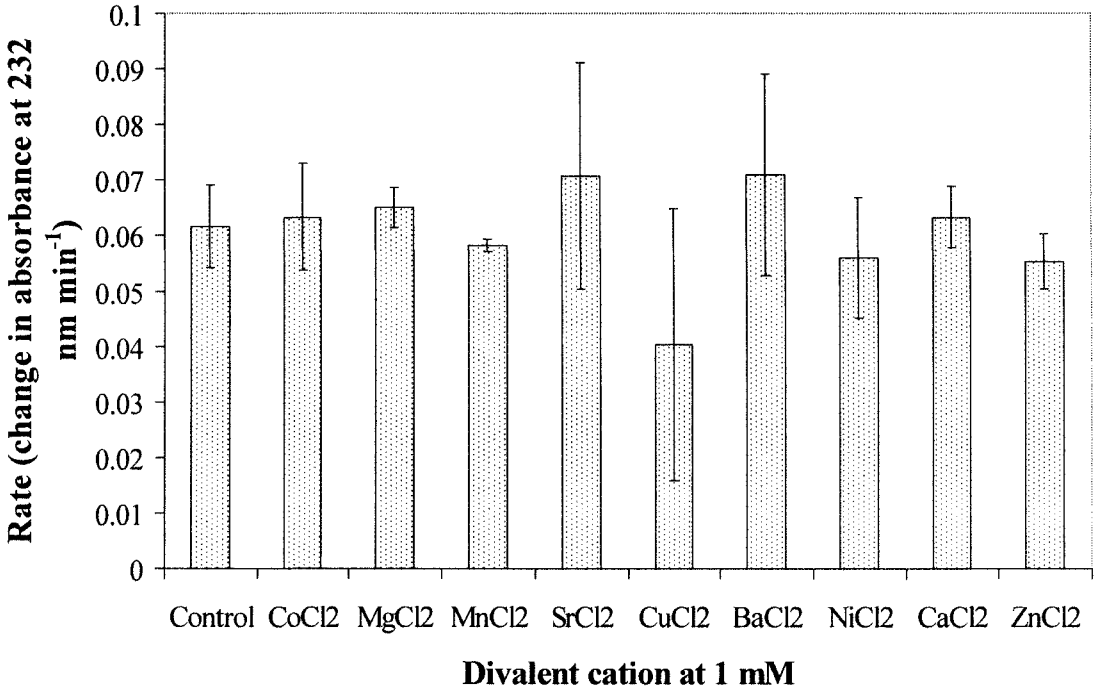


Figure 3.7: The effect of divalent cations on enzyme activity. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added and were performed by the method of Baker *et al*, (2002). Each divalent cation was added at a final concentration of 1 mM. Error bars represent the standard deviation from the mean.

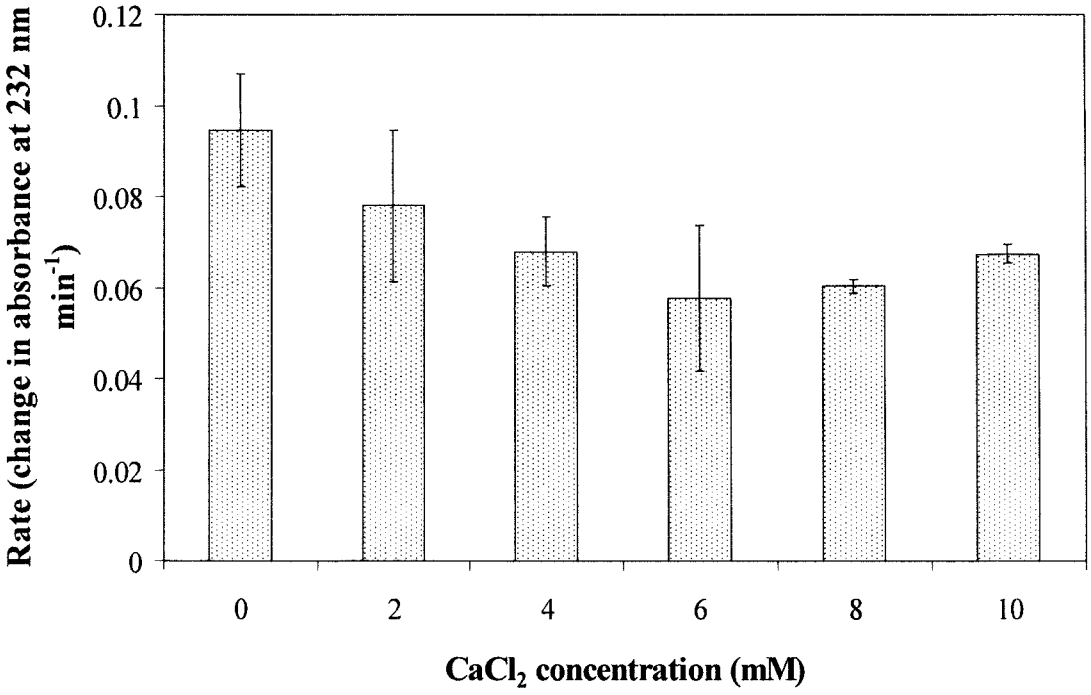


Figure 3.8: The effect of increased concentration of $CaCl_2$ on enzyme activity. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added and were performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

3.1.5 Site directed mutagenesis of HylP1

3.1.5.1 Alignment of HylP1 with a glycoside hydrolase from *Apis mellifera*

As it had previously been assigned to CAZy family GH69, an alignment of HylP1 was performed against a glycoside hydrolase from *Apis mellifera* using ClustalW. Catalytic residues determined in this enzyme were also observed in the sequence of HylP1 suggesting a role for these residues in the activity of HylP1 (figure 3.9). The location of these residues within the structure of HylP1 was seen to fall within a suspected area of the catalytic active site.

```
gi|13621887|gb|AAK33657.1|      -MSENIPLRVQFKRMKAAEWARSDVILLESEIGFETDTGFARAGDGHNRF 49
gi|155680|gb|AAA27730.1|        -MSRPLVITEGMMIGVLLMLAPINALLLGFVQSTPDNNKTVREFNVYWNV 49

gi|13621887|gb|AAK33657.1|      SDLGYIS----------------------------PLDYNLLTNKPNIDG 71
gi|155680|gb|AAA27730.1|        PTFMCHKYG----------------------------LRFEEVSEKYGILQ 72

gi|13621887|gb|AAK33657.1|      LATKVETAQKLQQ----KADKETVYTKAESKQELDKKLNLKGGVMTGQLK 117
gi|155680|gb|AAA27730.1|        NWMDKFRGEEIAILYDPGMFPALLKDPNGNVVARNGGVPQLGNLTKHLQV 122
                                                          :   *.  .

gi|13621887|gb|AAK33657.1|      FKPA-ATVAYSSSTGGAVNIDLSSTRGAGVVVYSDNDTSDGPLMSLRTGK 166
gi|155680|gb|AAA27730.1|        FRDHLINQIPDKSFPGVGVIDFESWRPIFRQNWASLQPYKKLSVEVVRRE 172
                                ::        ..*  *.  **:..        ::. :. .   : :   :

gi|13621887|gb|AAK33657.1†      ETFNQSALFVDYKG----TTNAVNIAMRQPTTPNFSSALNITSGNENGSA 212
gi|155680|gb|AAA27730.1|        HPFWDDQRVEQEAKRRFEKYGQLFMEETLKAAKRMRPAANWGYYAYPYCY 222
                                ..* ..  . :       . . : :     .  .: .* *        .

gi|13621887|gb|AAK33657.1|      MQLRGSEKALGTLKITHENPSIGADYDKNAAALSIDIVKK--TNGAGTAA 260
gi|155680|gb|AAA27730.1|        NLTPNQPSAQCEATTMQENDKMSWLFESEDVLLPSVYLRWNLTSGERVGL 272
                                    .       .  :** .    ::.: .  .              ..

gi|13621887|gb|AAK33657.1|      QGIYINSTSGTTGKLLRIRN------------LSDDKFYVKSDGGFYAK 297
gi|155680|gb|AAA27730.1|        VGGRVKEALRIARQMTTSRKKVLPYYWYKYQDRRDTDLSRADLEATLRKI 322

gi|13621887|gb|AAK33657.1|      ETSQIDGNLKLKDP----TANDHAATKAYVDKAISELKKLILKK------ 337
gi|155680|gb|AAA27730.1|        TDLGADGFIIWGSSDDINTKAKCLQFREYLNNELGPAVKRIALNNNANDR 372
```

```
gi|13621887|gb|AAK33657.1|          ----------
gi|155680|gb|AAA27730.1|            LTVDVSVDQV 38
```

Figure 3.9 ClustalW alignment of HylP1 (gi | 13621887 | gb | AAK33657.1 |) with a glycoside hydrolase from *Apis mellifera* (gi | 155680 | gb | AAA27730.1 |). Numbering used is not the same as used in the structure. Catalytic residues observed in gi | 155680 | gb | AAA27730.1 | which are also seen in gi | 13621887 | gb | AAK33657.1 | are highlighted in red.

3.1.5.2 Site directed mutagenesis of suspected catalytic residues

Following alignment of HylP1 with a glycoside hydrolase from A. mellifera, the location of these residues within the structure of HylP1 suggested the likely area of the catalytic active site. Tyrosine residues, known to play a role within the active site of polysaccharide lyases acting as a catalytic base, were also determined in this region and targeted for site directed mutagenesis. Residues D137, I136 and Y149 were successfully mutated and demonstrated a loss of catalytic activity against dialysed potassium hyaluronate as shown in figure 3.10.

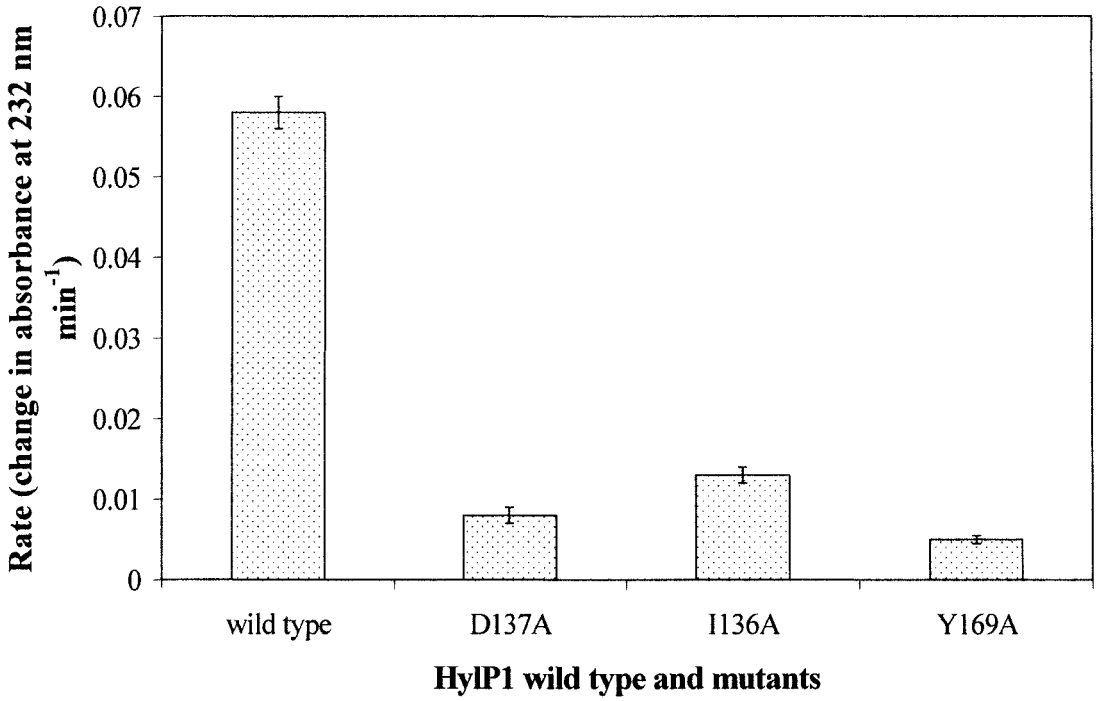


Figure 3.10: The effect of site directed mutation on enzyme activity. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added and were performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

3.1.5.3 Circular dichroism of native and mutated HylP1

To ascertain if the loss of catalytic activity in the mutated forms of HylP1 was the result of the loss of catalytic residues or a conformational change in the protein, circular dichroism was performed. Upon analysis it could be seen that D137A and Y149A were folded similarly to the native protein giving similar spectra, whereas the drastically different spectra seen in I136A is indicative of incorrect folding and hence a conformational change of the protein. Spectra can be seen in figure 3.11

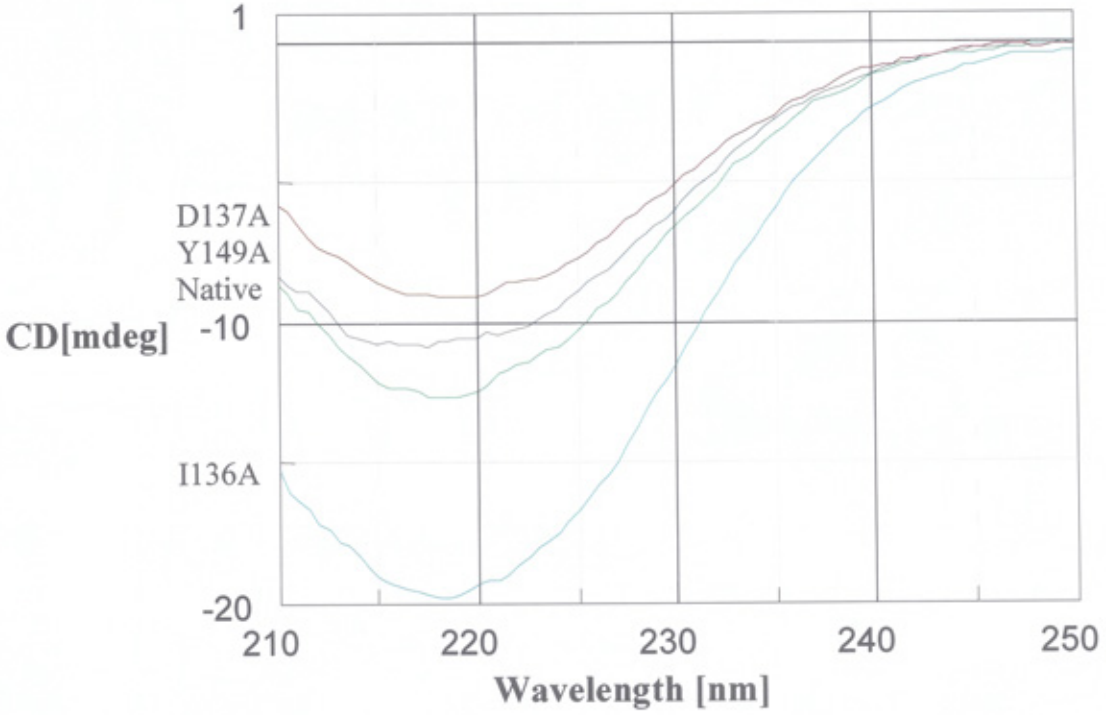


Figure 3.11: CD spectra of native HylP1 and the D137A, I136A and Y149A mutants. Data was collected as described in section 2.2.8.

3.1.6 Inhibition studies of HylP1

3.1.6.1 Inhibition of HylP1 by ascorbic acid

The effect of a known hyaluronate lyase inhibitor, ascorbic acid was determined by its inclusion in reactions at various concentrations. At all concentrations of ascorbic acid, HylP1 maintained activity as seen in figure 3.12 which indicates that the enzyme is not inhibited by ascorbic acid.

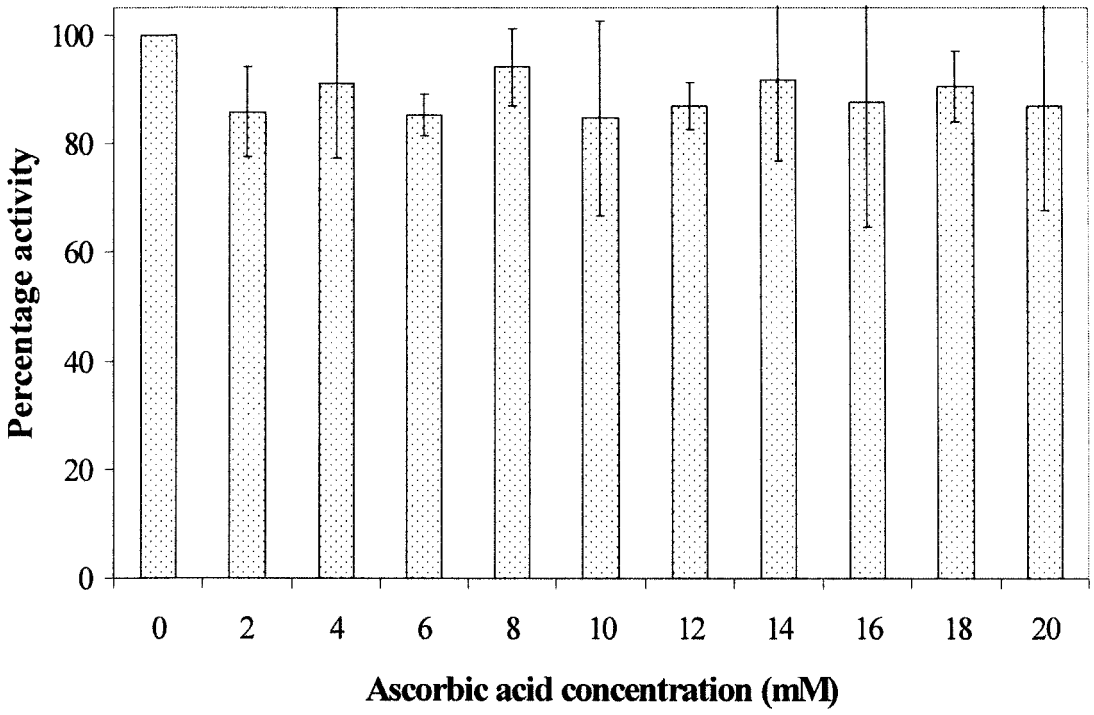
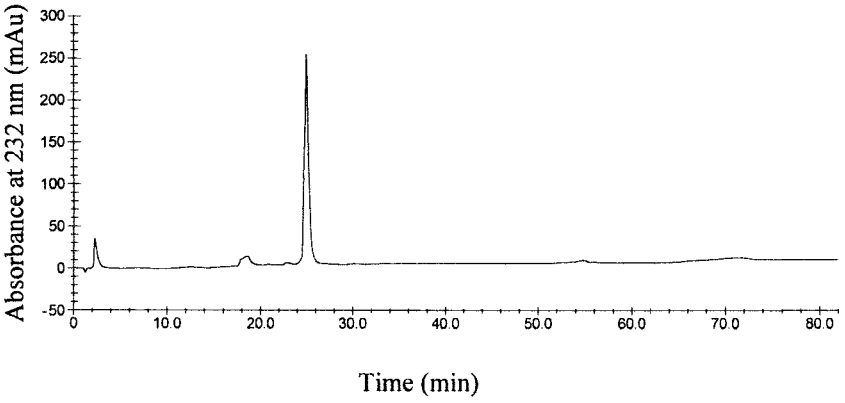


Figure 3.12: The effect of varying concentrations of ascorbic acid on enzyme activity. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added and were performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

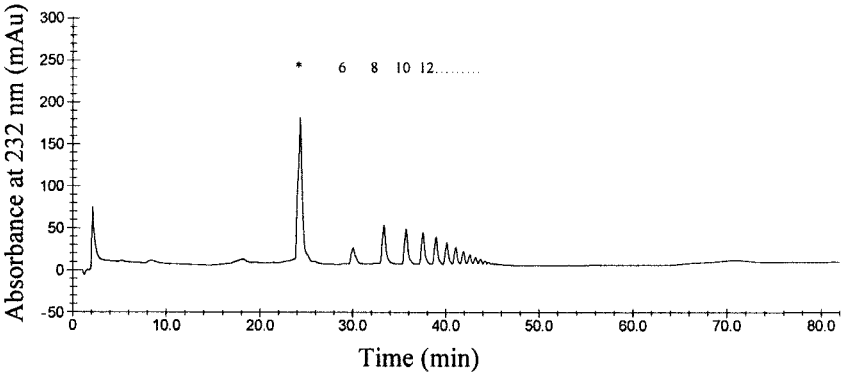
3.1.7 Mode of action studies of HylP1

To determine the mode of action and the size of products released, an enzyme reaction was set up as described in section 2.2.5.1.2. Aliquots were removed at time points, boiled to terminate the reaction and subjected to HPAEC analysis as described in section 2.2.7. Figure 3.13 shows HPAEC traces following digestion of dialysed potassium hyaluronate by HylP1. Upon analysis of the traces it is seen that HylP1 displays a random endolytic mode of action with a range of oligosaccharides produced, the smallest of which is a tetramer as determined by retention times.

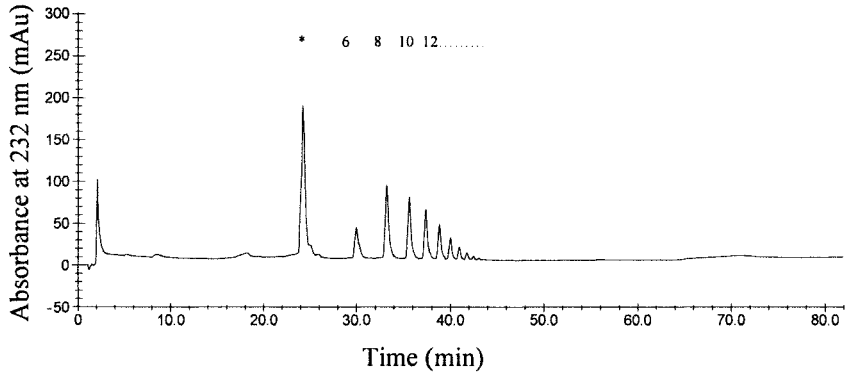
A


Absorbance at 232 nm (mAu)
300
250
200
150
100
50
0
-50
0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
Time (min)

B


Absorbance at 232 nm (mAu)
300
250
200
150
100
50
0
-50
*
6
8
10
12
0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
Time (min)

C



D

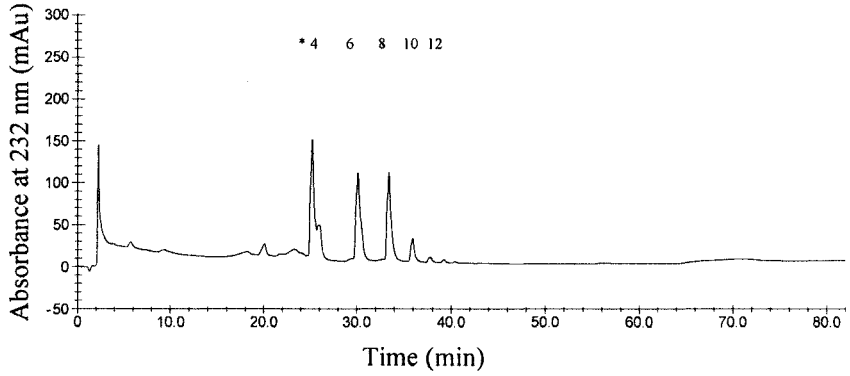


Figure 3.13. HPAEC analysis of HylP1 mode of action. Reaction time A: 0 minute; B: 10 minutes; C: 20 minutes; D: overnight digestion (* artifactual peak) Reactions were performed in triplicate with a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added. Digestion reactions were performed by the method of Baker *et al,* (2002) and HPAEC was performed according to the method of Lauder *et al*, (2000).

3.2 Triple stranded β helix of HylP1

3.2.1 Cloning, expression and purification of the triple stranded β helix of HylP1

The gene encoding the truncated form of HylP1, comprised of the central triple stranded β helix forming region, was amplified by PCR from the genomic DNA of *S. pyogenes* SF370. A primer pair (table 2.5) was used to generate a PCR product of the gene and this was ligated into PCR Blunt and digested out by an *Nde* I and *Xho* I. The fragment was cloned into a similarly restricted pET-28a vector to generate a plasmid encoding the protein with an N terminal His_6-tag (figure 3.14)

This was cloned into *E.coli* TOP10 cells from which the plasmid was extracted and transformed into *E. coli* BL21 cells allowing for expression and isolation of the protein as described in sections 2.2.2.2.3, 2.2.4.2 and 2.2.4.3. The protein fractions following IMAC were checked for purity by SDS PAGE (figure 3.15), concentrated and the concentration determined by Bradfords assay.

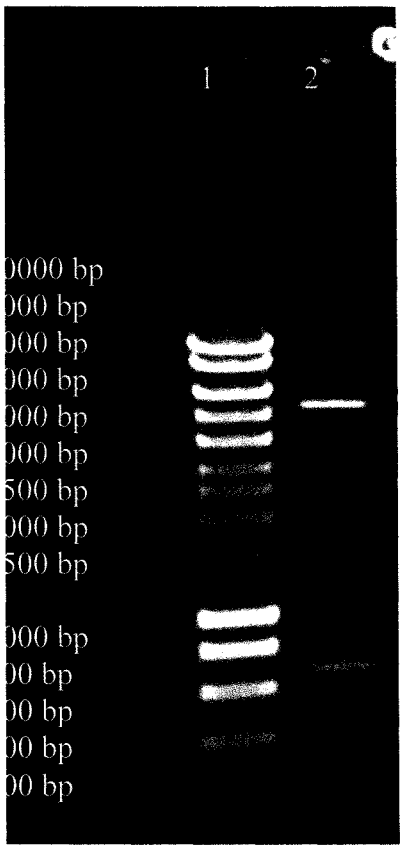


Figure 3.14: Visualisation of DNA bands following gel electrophoresis. Lane 1 Bioline Hyperladder I; lane 2 DNA fragments generated during a *Nde* I/*Xho* I restriction digest of pET-28a carrying the gene encoding the triple stranded β helix of HylP1.

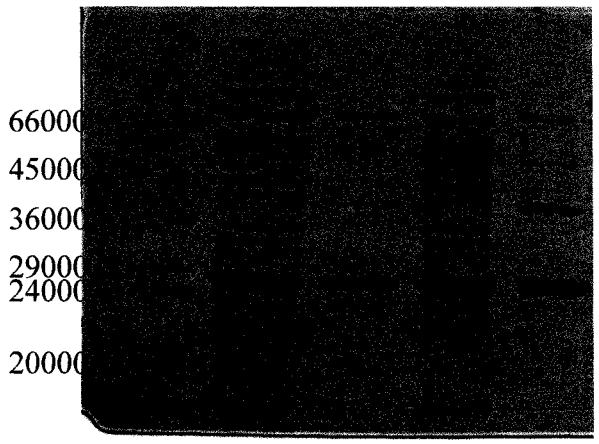


Figure 3.15: 12 % SDS PAGE of the triple stranded β helix of HylP1. Lane 1, Low molecular weight markers; Lane 2, triple stranded β helix cell free extract produced from cells grown at 20 °C; Lane 3, IMAC purified triple stranded β helix produced from cells grown at 20 °C; Lane 4, triple stranded β helix cell free extract produced from cells grown at 30 °C; Lane 5, IMAC purified triple stranded β helix produced from cells grown at 30 °C.

3.2.2 Kinetic analysis of the triple stranded β helix of HylP1

3.2.2.1 Michaelis-Menten parameters for the triple stranded β helix of HylP1

For the investigation of the activity of the triple stranded β helix HylP1 against hyaluronate as a lyase, the enzyme was assayed for an increase in absorbance at 232 nm against dialysed potassium hyaluronate. For absolute kinetic analysis, 4 to 5 substrate concentrations (determined empirically) were chosen and reactions performed in triplicate. To ascertain if the data reflected true Michaelis-Menten kinetics, a Lineweaver-Burk plot from the three replicates was constructed as seen in figure 3.16 and used for the determination of a K_m of 0.53 mg ml^{-1} and a K_{cat} of 11.1s^{-1}.

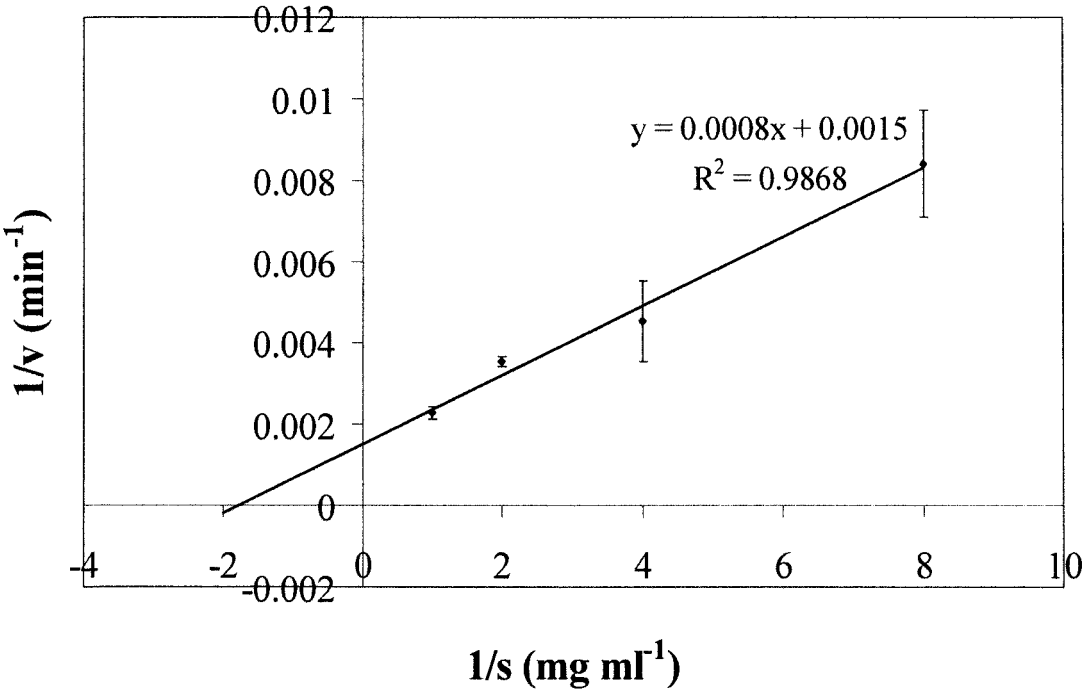


Figure 3.16: Lineweaver-Burke plot for the triple stranded β helix of HylP1. The reactions were performed in triplicate as detailed in section 2.2.6.1.2. Each reaction had 1 μg of enzyme added and was performed by the method of Baker *et al*, (2002). A linear trendline was added and error bars represent the standard deviation from the mean.

3.2.3 Biochemical and biophysical parameters of the triple stranded β helix of HylP1

3.2.3.1 Determination of pH optimum

The triple stranded β helix of HylP1 was shown to have an optimum pH of 6.5 against dialysed potassium hyaluronate. A range of buffers were used to determine the pH optimum. Assays were carried out at 37 °C and the rate measured spectrophotometrically at 232 nm. Figure 3.17 shows the effect of pH on enzyme activity with the curve being constructed from 3 independent data sets.

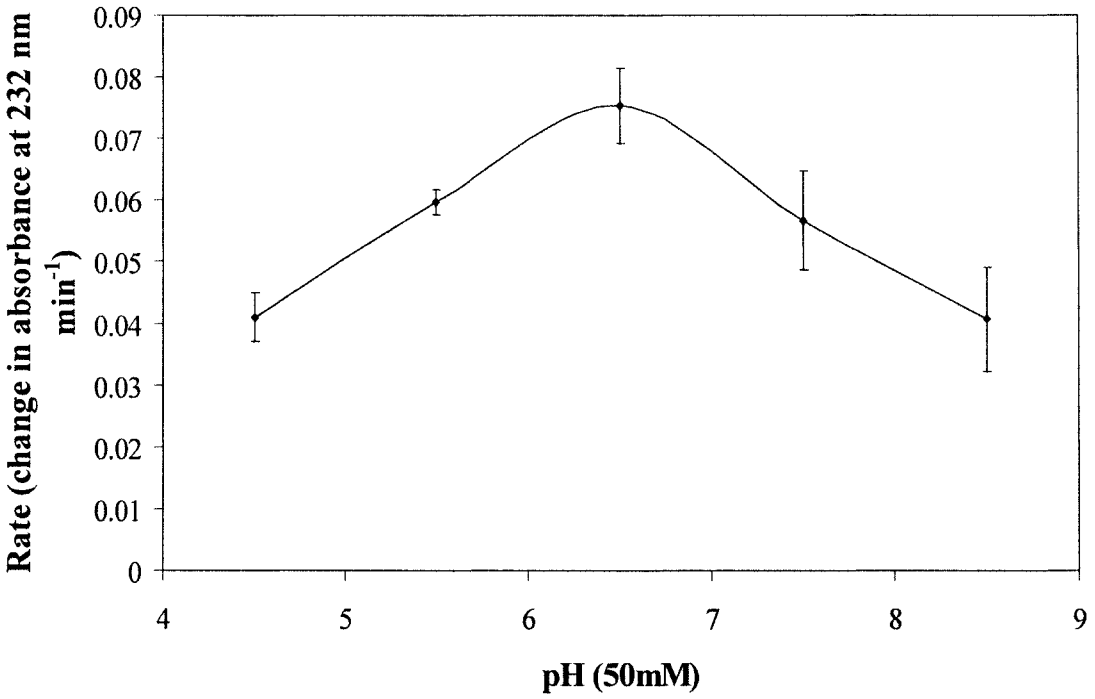


Figure 3.17: The effect of pH on the rate of enzyme activity against dialysed potassium hyaluronate. The reactions were performed in triplicate as detailed in section 2.2.6.1.2. Each reaction had a final substrate concentration of 2 mg ml^{-1} and 1 μg of enzyme added and was performed by the method of Baker *et al*, (2002). A polynomial trendline was added and error bars represent the standard deviation from the mean.

3.2.3.2 Determination of temperature optimum and thermostability

At 37 °C, HylP1 demonstrated the highest amount of activity against dialysed potassium hyaluronate when added chilled to a pre-warmed reaction mixture as seen in figure 3.18. At temperatures above and below 37 °C the enzyme exhibited reduced activity.

Upon pre-incubation for 20 min at various temperatures prior to assay, HylP1 retained highest activity at 27 °C and activity was significantly reduced at temperatures above 47 °C as seen in figure 3.19.

The effect of temperature on activity and stability was determined from 3 independent data sets.

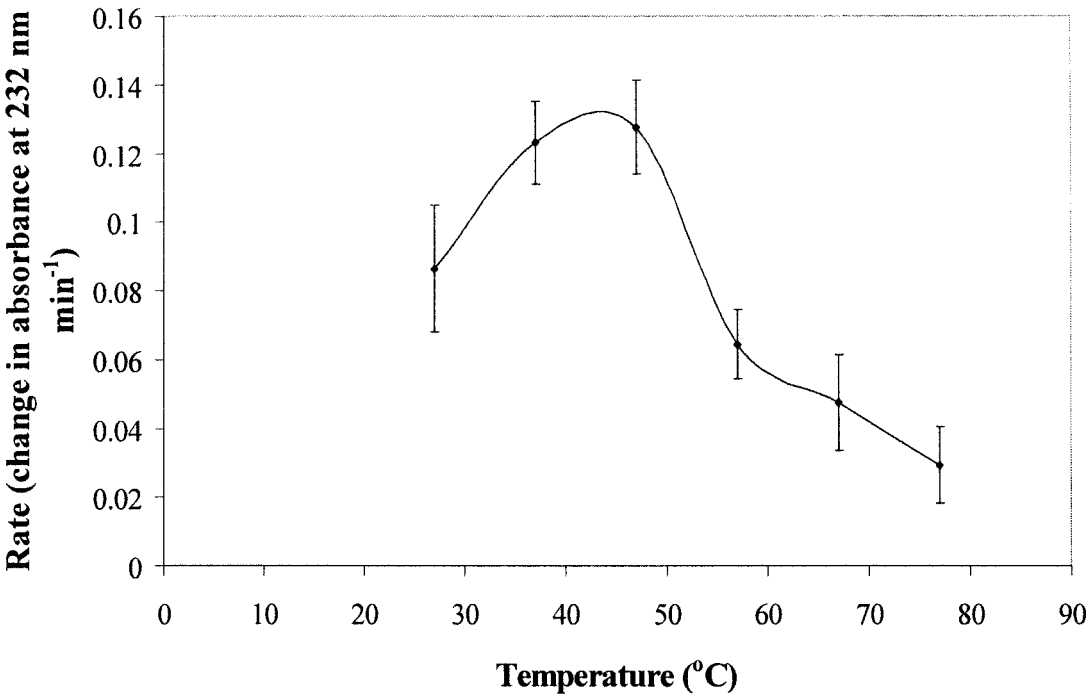


Figure 3.18: The effect of temperature on the rate of enzyme activity. The reactions were performed in triplicate as detailed in section 2.2.6.1.2. Each reaction had a final substrate concentration of 2 mg ml^{-1} and 1 μg of enzyme added and was performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

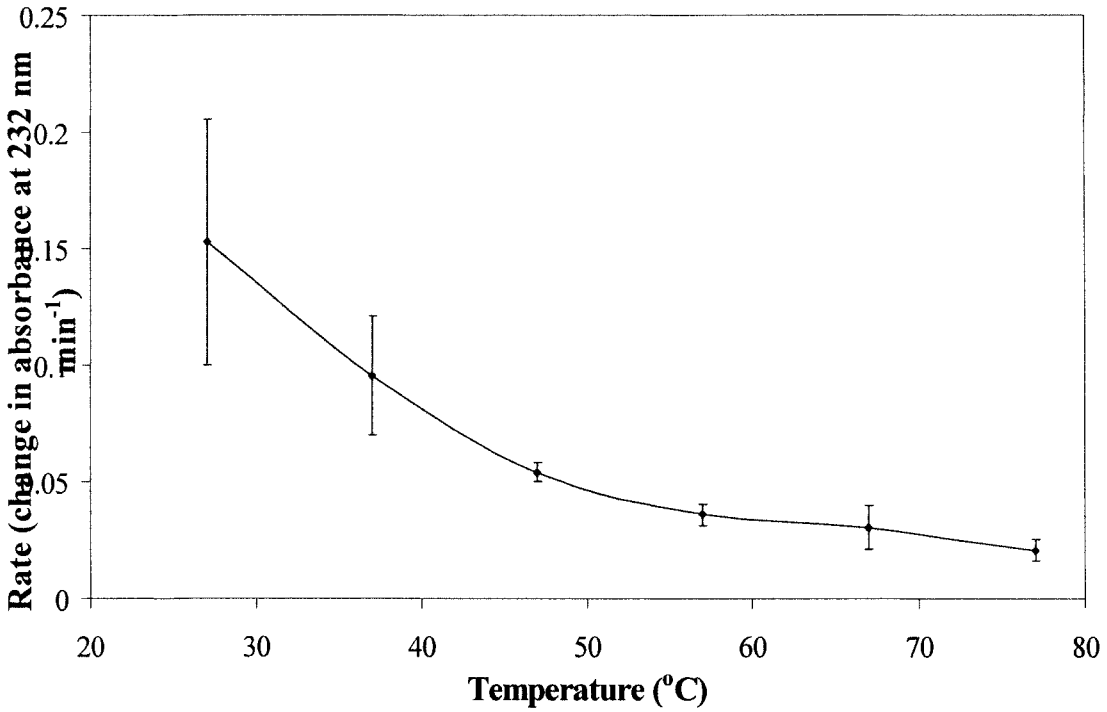


Figure 3.19: The effect of temperature on the stability of the enzyme. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} and 1 μg of enzyme added after pre incubation of the enzyme for 20 min at 27, 37, 47, 57, 67 and 77 °C and were performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

3.2.3.3 Determination of divalent ion requirement

When HylP1 was tested with a range of divalent cations, it was observed that the enzyme had absolutely no requirement for any of the ions tested. The data for the effect of divalent ions on activity is presented in figure 3.20. Activity was seen to decrease slightly when $CaCl_2$ and $CoCl_2$ were added and an increase was observed when $MnCl_2$ was added.

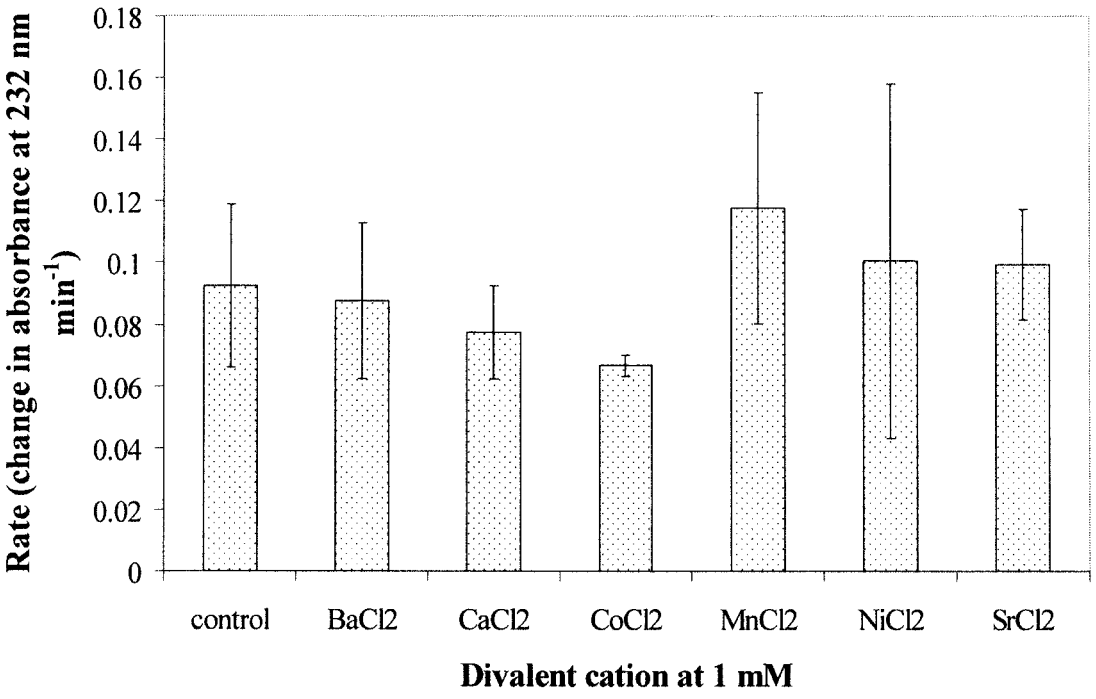


Figure 3.20: The effect of divalent cations on enzyme activity. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added and were performed by the method of Baker *et al*, (2002). Each divalent cation was added at a final concentration of 1 mM. Error bars represent the standard deviation from the mean.

3.3 SEQ2045

3.3.1 Alignment of HylP1 with the completed genome of *S. equi* 4047.

The recently completed genome of *S. equi* (available for searching at http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_equi) was BLAST searched with the sequence of HylP1 (accession number AAK33657). A sequence, SEQ2045, was observed to share 85 % identity with HylP1 as seen in figure 3.21 indicating a likely similar enzyme.

Score = 2751 (418.8 bits), Expect = 7.6e-120, P = 7.6e-120
Identities = 655/765 (85%), Positives = 655/765 (85%), Strand = Minus / Plus
[HSP Sequence]

```
Query:     989 AACTCAGAAATTGCTTTATCTACATAAGCTTTGGTTGCCGCATGATCATTCGCTGTGGGG 930
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2071749 AACTCAGAAATTGCTTTATCTACATAAGCTTTGGTTGCCGCATGATCATTCGCTGTGGGG 2071808

Query:     929 TCCTTGAGTTTCAGGTTGCCATCAATCTGCGAAGTTTCCTTGGCATAAAAACCACCGTCA 870
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2071809 TCCTTGAGTTTCAGGTTGCCATCAATCTGCGAAGTTTCCTTGGCATAAAAACCACCGTCA 2071868

Query:     869 GACTTGACGTAGAACTTATCATCACTAAGGTTTCTAATCCTAAGCAACTTCCCTGTCGTG 810
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2071869 GACTTGACGTAGAACTTATCATCACTAAGGTTTCTAATCCTAAGCAACTTCCCTGTCGTG 2071928

Query:     809 CCTGAGGTTGAGTTAATGTAGATTCCCTGAGCGGCTGTTCCTGCACCG--TTTGTCTTTT 752
               ||||| |||||||||||||||||| ||||| || || |||   |||  ||| | ||||
Sbjct: 2071929 CCTGATGTTGAGTTAATGTAGATTCCTTGAGCAGCAGTACCTTTTCCGCCTTTCTGTTTT 2071988

Query:     751 TGA-CAATATCAATGGATAACGCTGCCGCATTTTTATCATAATCCGCTCCAATACTTGGG 693
               | | | |||||||| |||||||||  || |||| ||| || |  ||  |||   |||||
Sbjct: 2071989 TTAACGATATCAATAGATAACGCTGTAGCGTTTTCATCGTATTTTGCCTCAACGTTTGGG 2072048

Query:     692 TTCTCATGAGTAATTTTTAGCGTTCCTAGCGCTTTTTCTGACCCTCGTAGCTGCATTGCA 633
               || || || || ||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2072049 TTTTCGTGTGTGATTTTGAGCGTTCCTAGCGCTTTTTCTGACCCTCGTAGCTGCATTGCA 2072108

Query:     632 CTACCATTTTCATTGCCGCTAGTAATATTAAGCGCCGATGAAAAATTGGGGGTGGTTGGC 573
               ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2072109 CTACCATTTTCATTGTCGCTAGTAATATTAAGCGCCGATGAAAAATTGGGGGTGGTTGGC 2072168
```

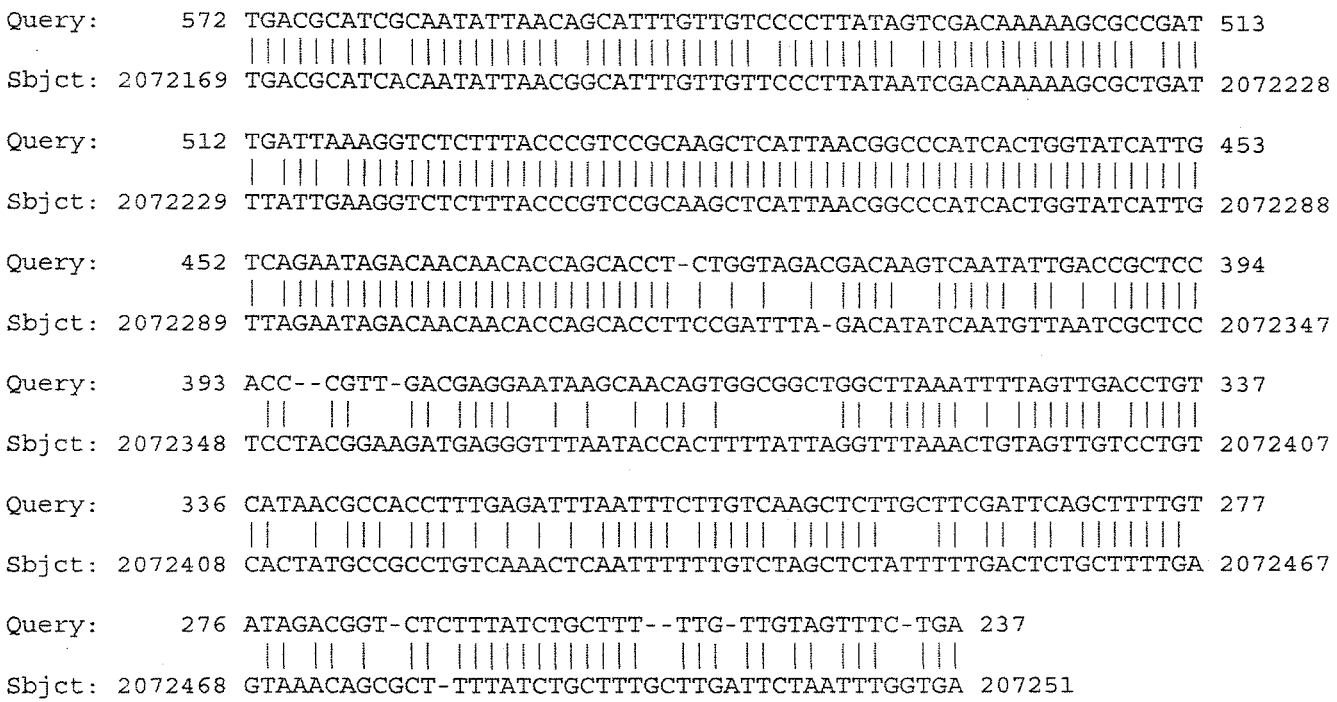


Figure 3.21: BLAST search result of HylP1 against the complete sequenced genome of *S. equi*

BLAST search was carried with HylP1 gene sequence out at http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_equi

```
SEQ2045         GPQGPSGERGLTGPIGPRGPAGKPGTTDYNQLQNKPNLDAFAQKKETNSKITKLESSKAD 60
HylP1           ----------------------------------------GLATKVETAQKLQ----QKAD 17
                                                        .:* * ** .*:     .***

SEQ2045         KSAVYSKAESKIELDKKLSLTGGIVTGQLQFKPNKSGIKPSSSVGGAINIDMSKSEGAGV 120
HylP1           KETVYTKAESKQELDKKLNLKGGVMTGQLKFKPAAT-VAYSSSTGGAVNIDLSSTRGAGV 76
                *.:**:***** ******.*.**::****:***  : :  ***.***:***:*.:.****

SEQ2045         VVYSNNDTSDGPLMSLRTGKETFNKSALFVDYKGTTNAVNIVMRQPTTPNFSSALNITSD 180
HylP1           VVYSDNDTSDGPLMSLRTGKETFNQSALFVDYKGTTNAVNIAMRQPTTPNFSSALNITSG 136
                ****:******************:****************.*****************.

SEQ2045         NENGSAMQLRGSEKALGTLKITHENPNVEAKYDENATALSIDIVKKQKGGKGTAAQGIYI 240
HylP1           NENGSAMQLRGSEKALGTLKITHENPSIGADYDKNAAALSIDIVKKTNG-AGTAAQGIYI 195
                *************************.: *.**:**:********* :*  *********

SEQ2045         NSTSGTTGKLLRIRNLSDDKFYVKSDGGFYAKETSQIDGNLKLKDPTANDHAATKAYVDK 300
HylP1           NSTSGTTGKLLRIRNLSDDKFYVKSDGGFYAKETSQIDGNLKLKDPTANDHAATKAYVDK 255
                ************************************************************

SEQ2045         AISELKKLILKK 312
HylP1           AISELKKLILKK 267
                ************
```

Figure 3.22 ClustalW alignment of SEQ2045 with HylP1 (AAK33657). Clustal W was performed at http://www.ebi.ac.uk/Tools/clustalw/index.html

3.3.2 Cloning, expression and purification of the triple stranded β helix of HylP1

The gene encoding SEQ2045, was amplified by PCR from the genomic DNA of *S. equi* 4047. A primer pair (table 2.5) was used to generate a PCR product of the gene and this was ligated into PCR Blunt and digested out by an *Nde I* and *Bam HI*. The fragment was cloned into a similarly restricted pET-28a vector to generate a plasmid encoding the protein with an N terminal His_6-tag (figure 3.22)

This was cloned into *E.coli* TOP10 cells from which the plasmid was extracted and transformed into *E. coli* BL21 cells allowing for expression and isolation of the protein as described in sections 2.2.2.2.3, 2.2.4.2 and 2.2.4.3. The protein fractions following IMAC were checked for purity by SDS PAGE (figure 3.23), concentrated and the concentration determined by Bradfords assay.

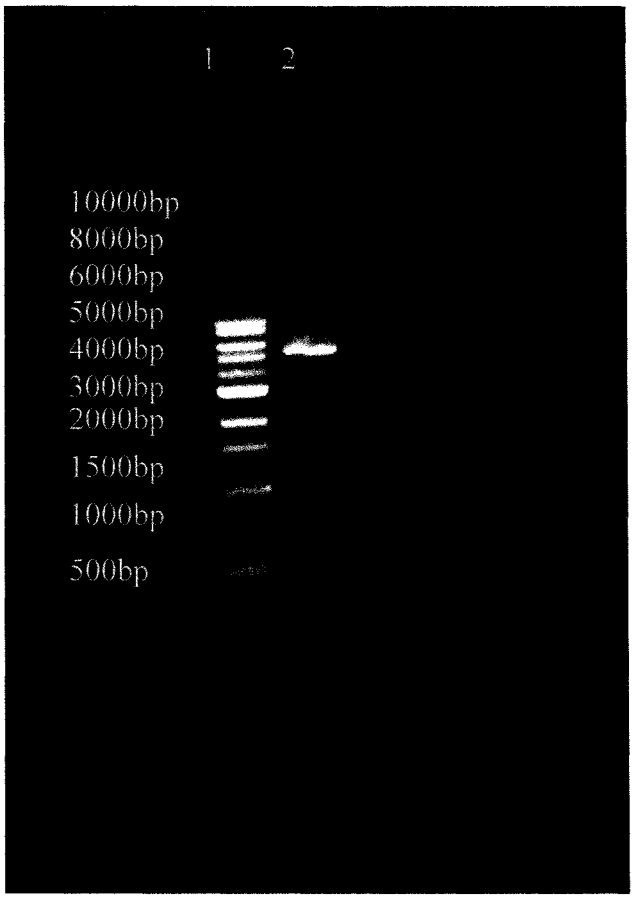


Figure 3.23: Visualisation of DNA bands following agarose gel electrophoresis. Lane 1: 1 μg DNA ladder (sizes 10, 8, 6, 5, 4, 3, 2, 1.5, 1, 0.5 kb). Lane 2: DNA fragments produced during a restriction endonuclease digest of cloned SEQ2045 in pET 28a vector. Approximately 0.3 μg of DNA was digested with *Nde* I and *Bam* HI (section 2.2.3.1.3)

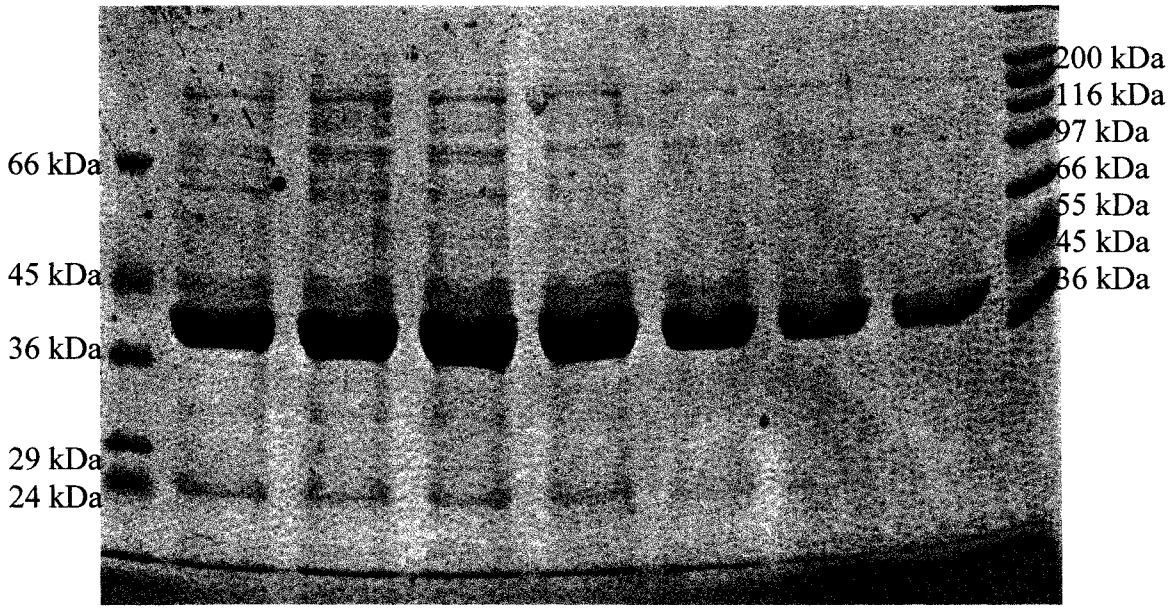


Figure 3.24: 12% (w/v) SDS PAGE of IMAC purified fractions of SEQ2045.

Lane 1: low molecular weight markers, lanes 2-8 IMAC purified fractions of SEQ2045, lane 9 high molecular weight markers. SEQ2045 was purified from *E. coli* cells harbouring the plasmid encoding the protein. A gradient elution technique as described in section 2.2.4.3.1 was employed for the purification of the protein from the cell free extract.

3.3.3 Kinetic analysis of SEQ2045

3.3.3.1 Michaelis-Menten parameters for SEQ2045

For the investigation of the activity of SEQ2045 against hyaluronate as a lyase, the enzyme was assayed for an increase in absorbance at 232 nm against dialysed potassium hyaluronate. For absolute kinetic analysis, 4 to 5 substrate concentrations (determined empirically) were chosen and reactions performed in triplicate. To ascertain if the data reflected true Michaelis-Menten kinetics, a Lineweaver-Burk plot from the three replicates was constructed as seen in figure 3.24 and used for the determination of a K_m of 0.33 mg ml^{-1} and a K_{cat} of 7.61 s^{-1}.

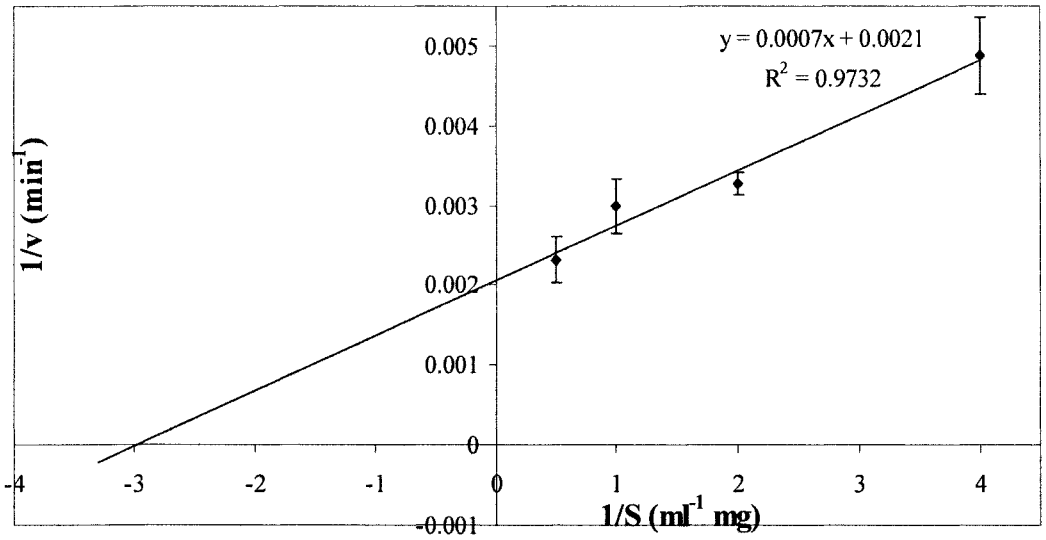


Figure 3.25: Lineweaver-Burke plot for SEQ2045 against dialysed potassium hyaluronate. The reactions were performed in triplicate with error bars representing the standard deviation fom the mean. The reactions were performed under standard assay conditions.

3.3.4 Biochemical and biophysical parameters of SEQ2045

3.3.4.1 Determination of pH optimum

SEQ2045 was shown to have an optimum pH of 6.5 against dialysed potassium hyaluronate. A range of buffers were used to determine the pH optimum. Assays were carried out at 37 °C and the rate measured spectrophotometrically at 232 nm. Figure 3.25 shows the effect of pH on enzyme activity with the curve being constructed from 3 independent data sets.

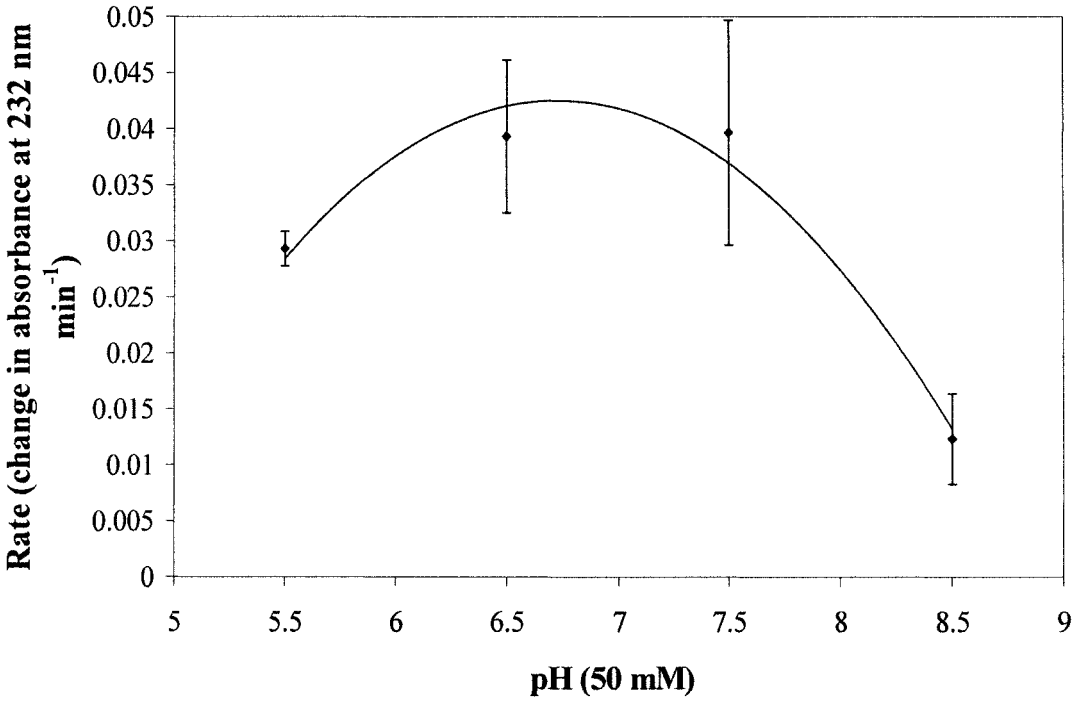


Figure 3.26: The effect of pH on the rate of enzyme activity against dialysed potassium hyaluronate. The reactions were performed in triplicate as detailed in section 2.2.6.1.2. Each reaction had a final substrate concentration of 2 mg ml^{-1} and 1 µg of enzyme added and was performed by the method of Baker *et al*, (2002). A polynomial trendline was added and error bars represent the standard deviation from the mean.

3.3.4.2 Determination of temperature optimum and thermostability

At 37 °C, SEQ2045 demonstrated the highest amount of activity against dialysed potassium hyaluronate when added chilled to a pre-warmed reaction mixture as seen in figure 3.26. At temperatures above and below 37 °C the enzyme exhibited reduced activity.

Upon pre-incubation for 20 min at various temperatures prior to assay, SEQ2045 retained highest activity at 27 °C and activity was significantly reduced at temperatures above 47 °C as seen in figure 3.27.

The effect of temperature on activity and stability was determined from 3 independent data sets.

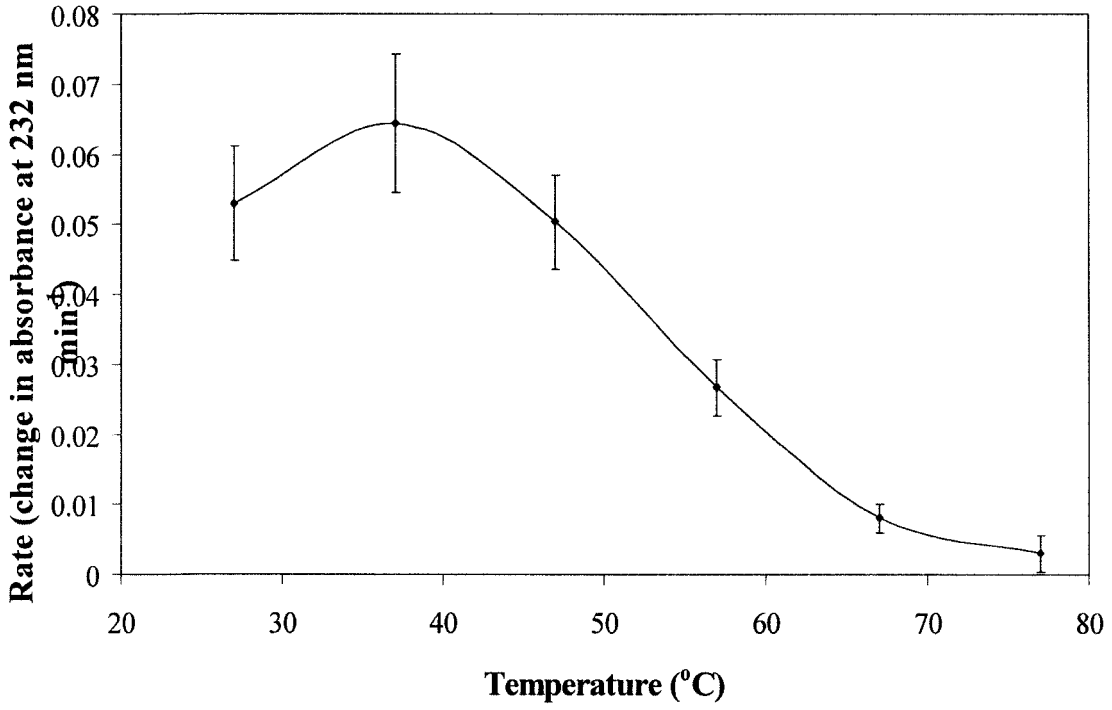


Figure 3.27: The effect of temperature on the rate of enzyme activity against dialysed potassium hyaluronate. The reactions were performed in triplicate as detailed in section 2.2.6.1.2. Each reaction had a final substrate concentration of 2 mg ml^{-1} and 1 μg of enzyme added and was performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

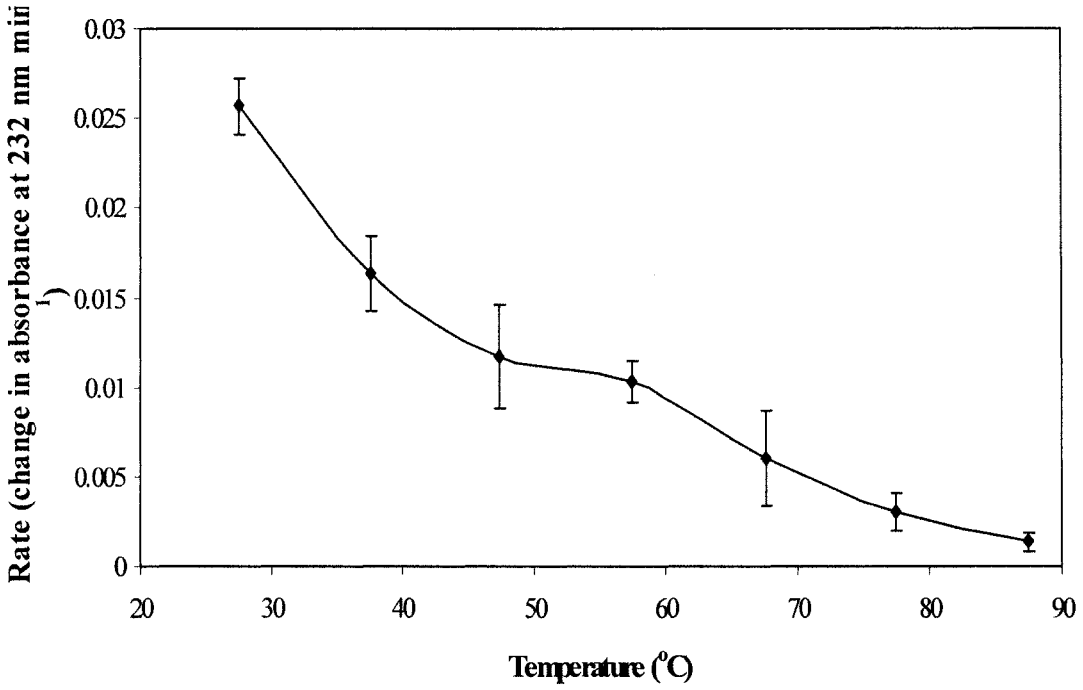


Figure 3.28: The effect of temperature on the stability of SEQ2045. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} and 1 μg of enzyme added after pre incubation of the enzyme for 20 min at 27, 37, 47, 57, 67 and 77 °C and were performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

3.3.4.3 Determination of divalent ion requirement

When SEQ2045 was tested with a range of divalent cations, it was observed that the enzyme had absolutely no requirement for any of the ions tested. The data for the effect of divalent ions on activity is presented in figure 3.28. Activity was seen to increase when $BaCl_2$, $CoCl_2$, $MgCl_2$ and $MnCl_2$.

A range of increased concentrations of CaCl2 were also added and SEQ2045 showed a requirement for $CaCl_2$ at 8mM as seen in figure 3.29.

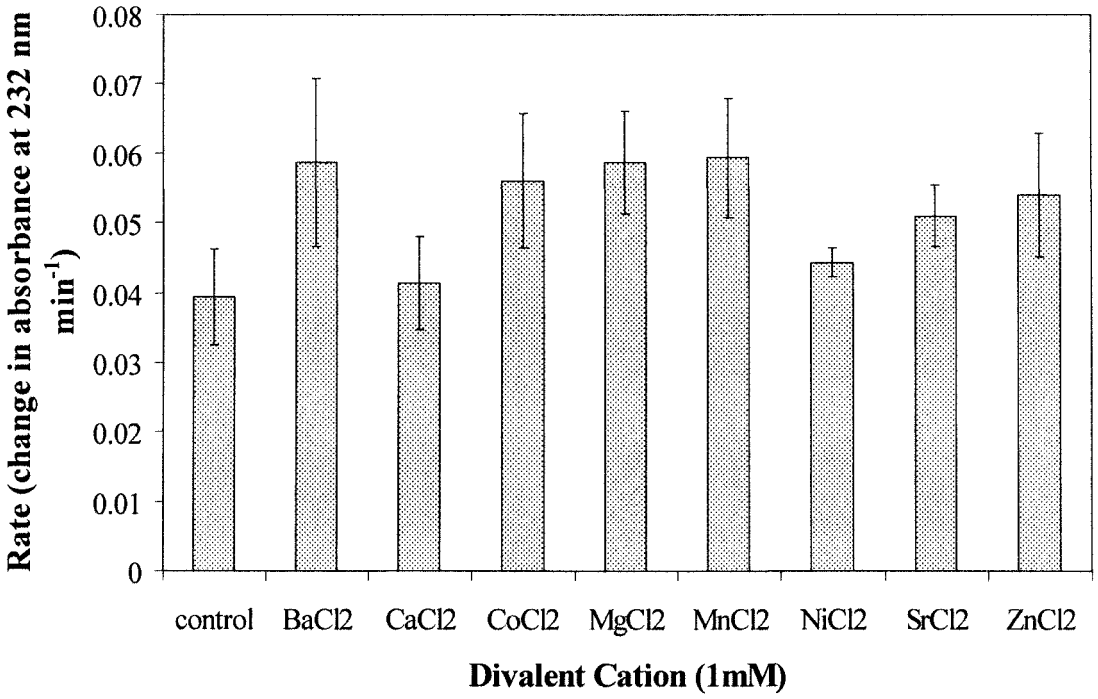


Figure 3.29: The effect of divalent cations on enzyme activity. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added and were performed by the method of Baker *et al*, (2002). Each divalent cation was added at a final concentration of 1 mM. Error bars represent the standard deviation from the mean.

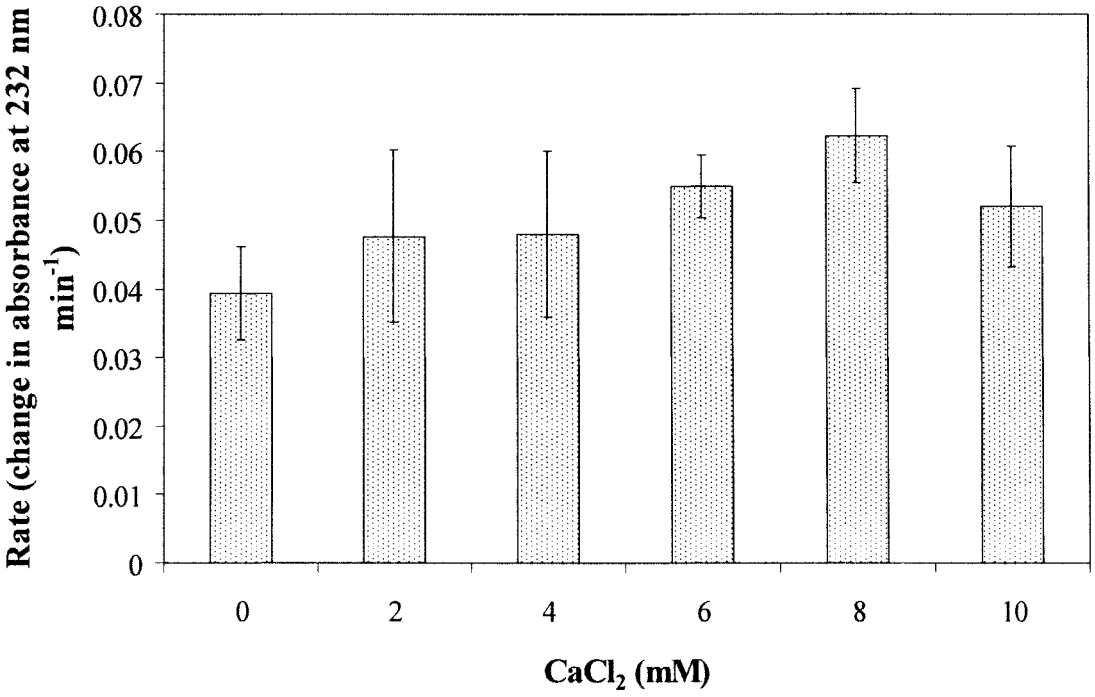


Figure 3.30: The effect of increased concentration of $CaCl_2$ on enzyme activity. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added and were performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

3.3.5 Site directed mutagenesis of SEQ2045

3.3.5.1 Alignment with HylP1

The sequence of HylP1 and SEQ2045 were aligned using ClustalW. Residues which had been observed to play a role in catalysis of HylP1 were also observed in SEQ2045 suggesting they were conserved residues important in the catalytic behaviour of the enzyme (figure 3.30).

```
CLUSTAL W (1.83) multiple sequence alignment

gi|13621887|gb|AAK33657.1|      MSENIP-LRVQFKRMKAAEWARSDVILLESEIGFETDTGFARAGDGHNRF 49
SEQ2045                         MSKEVASARIQHRGMTKQEWESSSDILMEREIGIDMTTGYPKVGDGKNKF 50
                                **:::.  *:*.: *.  **  *. **:* ***::  **:.:.***:*:*

gi|13621887|gb|AAK33657.1|      SDL-------------------GYISP---------LDYNLLTNKPNIDG 71
SEQ2045                         KDLKDLRGPMGPQGPSGERGLTGPIGPRGPAGKPGTTDYNQLQNKPNLDA 100
                                .**                    * *.*          *** * ****:*.

gi|13621887|gb|AAK33657.1|      LATKVETAQKLQ----QKADKETVYTKAESKQELDKKLNLKGGVMTGQLK 117
SEQ2045                         FAQKKETNSKITKLESSKADKSAVYSKAESKIELDKKLSLTGGIVTGQLQ 150
                                :* * ** .*:     .****.:**:***** ******.*.**::****:

gi|13621887|gb|AAK33657.1|      FKPAAT-VAYSSSTGGAVNIDLSSTRGAGVVVYSDNDTSDGPLMSLRTGK 166
SEQ2045                         FKPNKSGIKPSSSVGGAINIDMSKSEGAGVVVYSNNDTSDGPLMSLRTGK 200
                                ***  : :  ***.***:***:*.:.********:*************

gi|13621887|gb|AAK33657.1|      ETFNQSALFVDYKGTTNAVNIAMRQPTTPNFSSALNITSGNENGSAMQLR 216
SEQ2045                         ETFNKSALFVDYKGTTNAVNIVMRQPTTPNFSSALNITSDNENGSAMQLR 250
                                ****:***************.*****************.**********

gi|13621887|gb|AAK33657.1|      GSEKALGTLKITHENPSIGADYDKNAAALSIDIVKKTNG-AGTAAQGIYI 265
SEQ2045                         GSEKALGTLKITHENPNVEAKYDENATALSIDIVKKQKGGKGTAAQGIYI 300
                                ****************.: *.**:**:********* :*  *********



gi|13621887|gb|AAK33657.1|      NSTSGTTGKLLRIRNLSDDKFYVKSDGGFYAKETSQIDGNLKLKDPTAND 315
SEQ2045                         NSTSGTTGKLLRIRNLSDDKFYVKSDGGFYAKETSQIDGNLKLKDPTAND 350
                                **************************************************
```

```
gi|13621887|gb|AAK33657.1|      HAATKAYVDKAISELKKLILKK 337
SEQ2045                         HAATKAYVDKAISELKKLILKK 372
                                **********************
```

Figure 3.31: ClustalW alignment of HylP1 (gi | 13621887 | gb | AAK.33657.1) and SEQ2045. Red indicates the reserved residues shown to be involved in catalytis activity of HylP1.

3.3.5.2 Site directed mutagenesis of suspected catalytic residues

Following alignment of SEQ2045 and HylP1 suspected catalytic residues were targeted for site directed mutatgenesis. The subsequent proteins demonstrated a loss of catalytic activity as seen in figure 3.31.

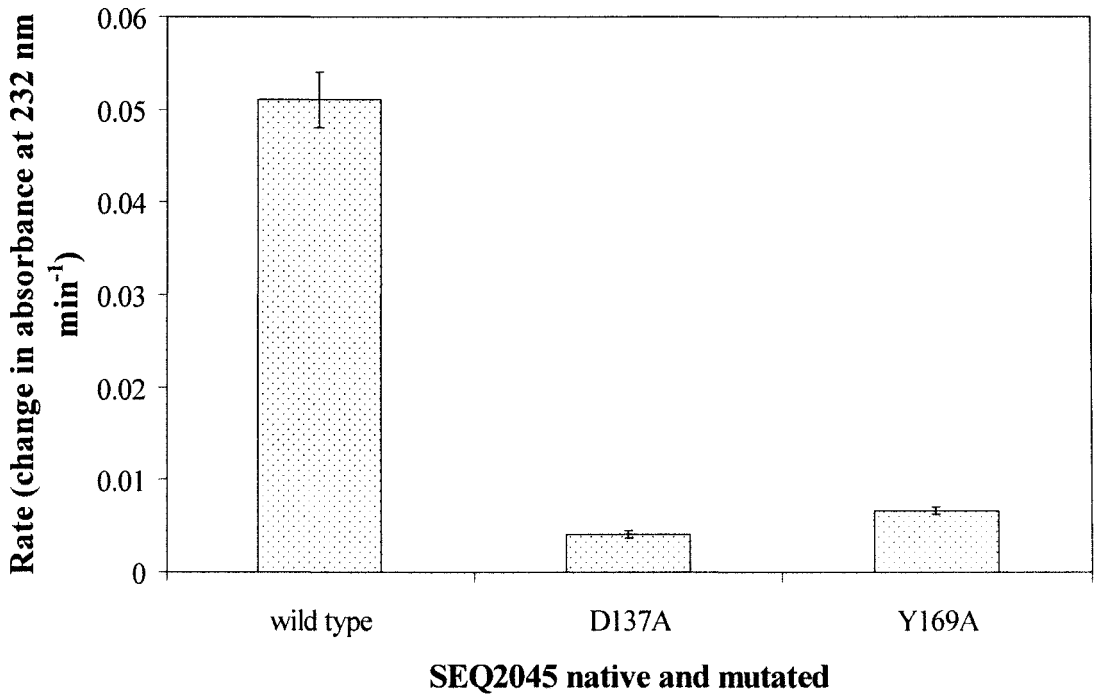


Figure 3.32: The effect of site directed mutation on enzyme activity. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added and were performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

3.3.6 Inhibition studies of SEQ2045

3.3.6.1 Inhibition of SEQ2045 by ascorbic acid

The effect of a known hyaluronate lyase inhibitor, ascorbic acid was determined by it's inclusion in reactions at various concentrations. At all concentrations of ascorbic acid, SEQ2045 maintained activity as seen in figure 3.32 which indicates that the enzyme is not inhibited by ascorbic acid.

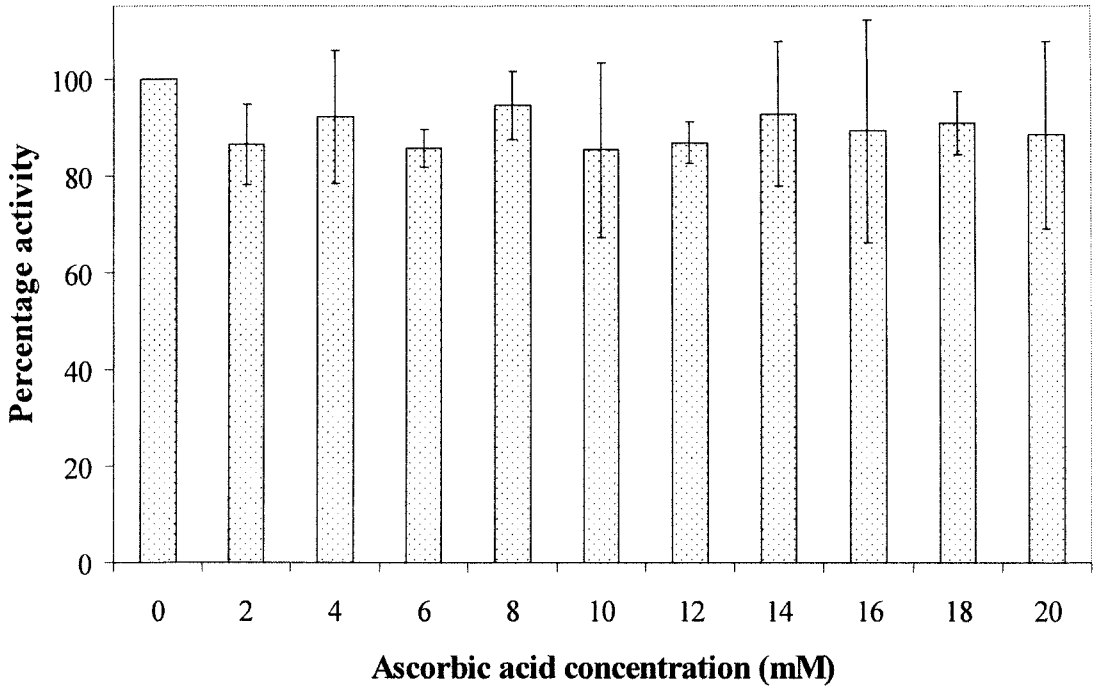
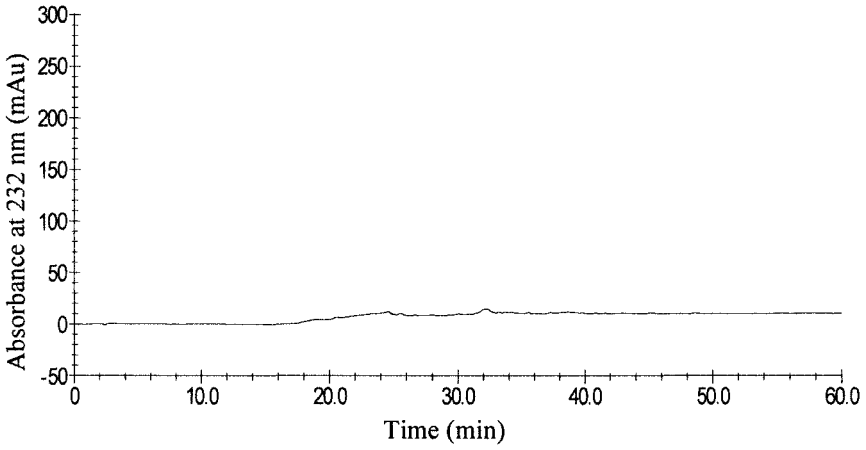


Figure 3.33: The effect of varying concentrations of ascorbic acid on enzyme activity. The reactions were performed in triplicate at a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added and were performed by the method of Baker *et al*, (2002). Error bars represent the standard deviation from the mean.

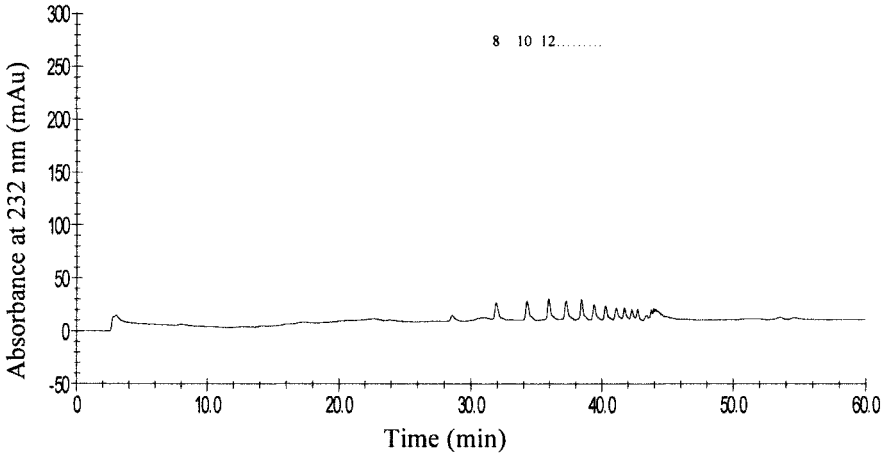
3.3.7 Mode of action studies of SEQ2045

To determine the mode of action and the size of products released, an enzyme reaction was set up as described in section 2.2.6.1.2. Aliquots were removed at time points, boiled to terminate the reaction and subjected th HPAEC analysis as described in section 2.2.7. Figure 3.33 shows HPAEC traces following digestion of dialysed potassium hyaluronate by SEQ2045. Upon analysis of the traces it is seen that SEQ2045 displays a random endolytic mode of action with a range of oligosaccharides produced, the smallest of which is a tetramer as determined by retention times.

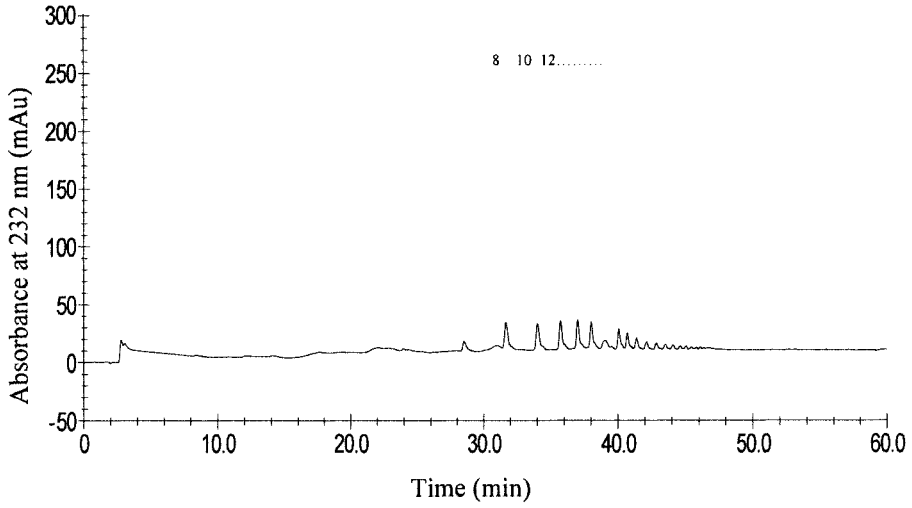
A


Absorbance at 232 nm (mAu)
300
250
200
150
100
50
0
-50
0
10.0
20.0
30.0
40.0
50.0
60.0
Time (min)

B


8 10 12
Absorbance at 232 nm (mAu)
300
250
200
150
100
50
0
-50
0
10.0
20.0
30.0
40.0
50.0
60.0
Time (min)

C



D

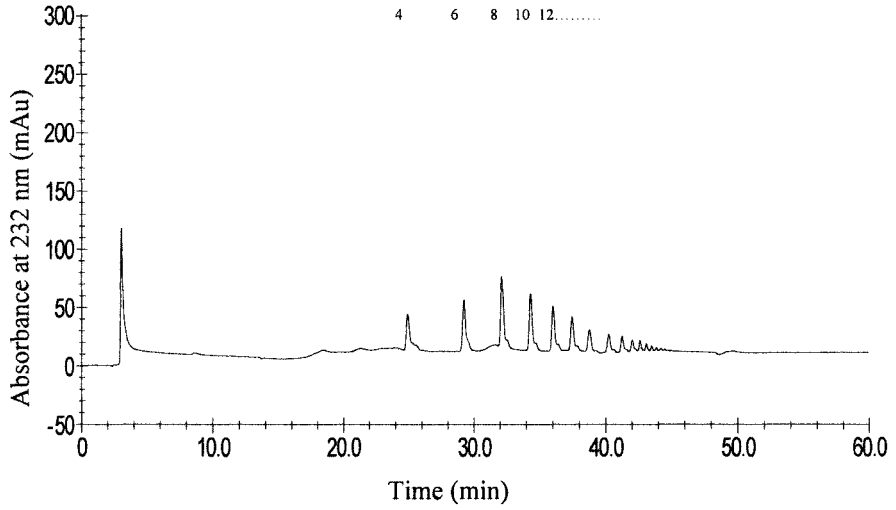


Figure 3.34: HPAEC analysis of SEQ2045 mode of action. Reaction time A: 0 minute; B: 10 minutes; C: 20 minutes; D: overnight digestion. Reactions were performed in triplicate with a final substrate concentration of 2 mg ml^{-1} with 1 μg of enzyme added. Digestion reactions were performed by the method of Baker *et al*, (2002) and HPAEC was performed according to the method of Lauder *et al*, (2000).

3.3.8 Western blot analysis of SEQ2045

To investigate the role of SEQ2045 as a potential virulence factor in *S. equi,* western blots were performed using serum taken from horses following infections. These sera were used as primary antibody and rabbit anti-horse antibody was used as a secondary antibody. Serum taken from a horse that has had no *S. equi* infection (to act as a negative control), an *S. equi* convalescent serum and an *R. equi* convalescent serum (to show the reaction is a specific result of a S. equi infection instead of a reaction simply as a esult of a non specific immune response) were compared against one another showing a reaction only in the serum taken following an *S. equi* infection as shown in figure 3.34. Figure 3.35 shows pre and post infection matched serum and again a clear reaction is seen following an *S. equi* infection.

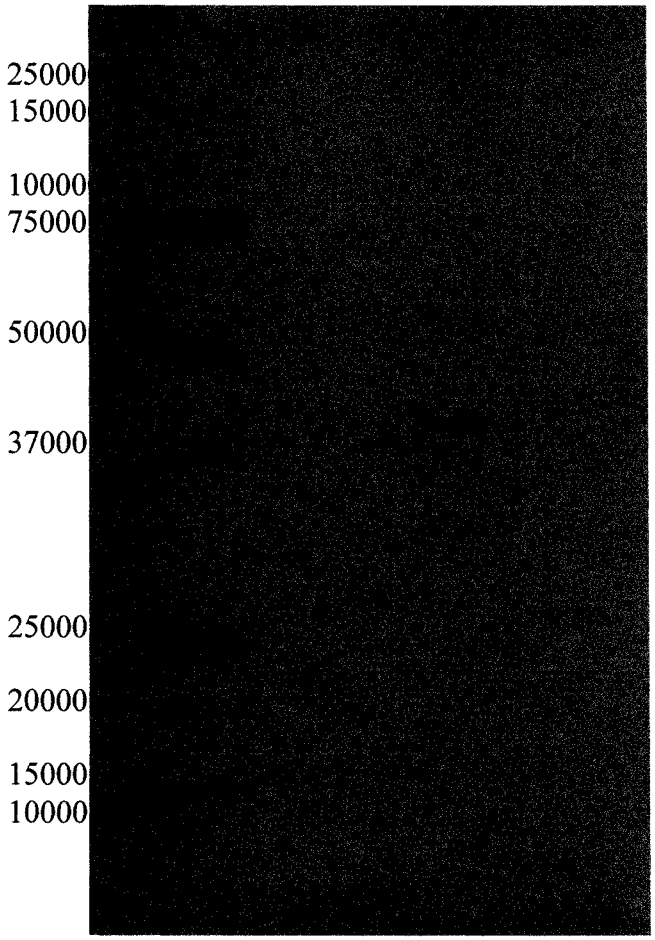


Figure 3.35: Western blot analysis of SEQ2045 with serums of different origins. Lane A; size standard (sizes 250, 150, 100, 75, 50, 37, 25, 20, 15, 10 kDa), Lane B; normal horse serum, Lane C; *S. equi* infected horse serum, Lane D; *R. equi* infected horse serum.

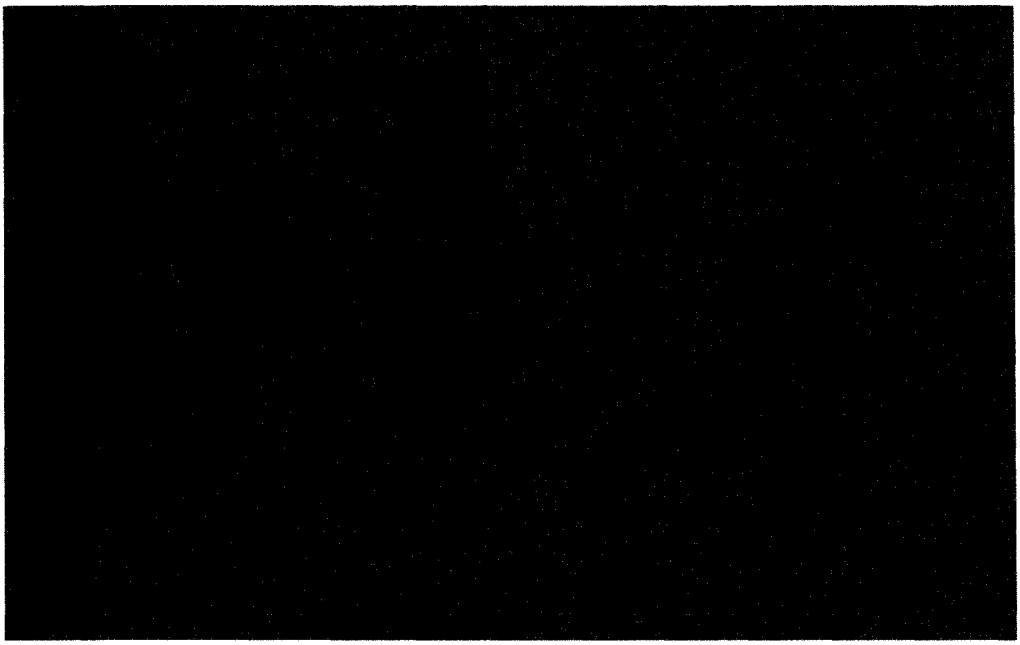


Figure 3.36: Western blot analysis of SEQ2045 with pre and post *S. equi* infection matched serums. Lane A; pony 1 pre infection, Lane B; pony 1 post infection, Lane C; size standard (sizes 250, 150, 100, 75, 50, 37, 25, 20, 15, 10 kDa), Lane D; pony 2 pre infection, Lane E; pony 2 post infection, Lane F; pony 3 pre infection, Lane G; pony 3 post infection.

3.3.9 Bioinformatic analysis of *S. equi*

SEQ2045 and HylP1 share 85 % identity so to ascertain if SEQ2045 is also phage encoded, bioinformatic analysis was performed using BLASTp searches. The putative gene sequences either side of SEQ2045 were BLASTp searched and the best match recorded. This was also performed by BLAST searching against M1 S. pyogenes SF370 to ascertain if the phage on which SEQ2045 was thought to lie may be the same phage as seen in *S. pyogenes*. Table 3.1 shows this data and indicates that SEQ2045 falls within a prophage region similar to HylP1. This data is further shown in diagrammatically in figure 3.36 where the prophage map of *S. equi* is compared with the prophage map of *S. pyogenes* SF370.1 showing distinct similarities between the two.

ORF[1]	Nucleotide Start codon	Nucleotide Stop codon	Best blastp match	E value of best match	Best blastp M1 GAS SF370 prophage match	E value of best M1 GAS SF370 prophage match	ORF/ Prophage
1 (SEQ2089)	2104394	2105533	Integrase [*Streptococcus pyogenes* M1 GAS]	0	see Best blastp match	see E value of best match	SPy0937/ SF370.2
2 (SEQ2088)	2104000	2104266	phage protein [*Streptococcus pyogenes* phage 5005.2]	5.00E-29	none	N/A	N/A
3 (SEQ2087)	2103602	2103988	similar to cI-like repressor, metallo-proteinase motif [*Streptococcus thermophilus* bacteriophage Sfi21]	2.00E-12	none	N/A	N/A
4 (SEQ2086)	2103249	2103599	transcriptional regulator, Cro/cI family [*Streptococcus agalactiae* H36B]	2.00E-36	putative repressor - phage associated	5.00E-05	SPy1486/ SF370.3
5 (SEQ2085)	2102953	2102708	hypothetical protein SAI_0575 [*Streptococcus agalactiae* H36B]	2.00E-24	none	N/A	N/A
6 (SEQ2084)	2101969	2102757	hypothetical protein SAI_0576 [*Streptococcus agalactiae* H36B]	3.00E-137	none	N/A	N/A
7 (SEQ2083)	2101918	2101727	Phage antirepressor [Bacteriophage Sal1]	3.00E-07	none	N/A	N/A

8 (SEQ2082)	2101646	2101335	prophage ps2 protein 07, excisionase [*Streptococcus agalactiae* H36B]	1.00E-45	putative excisionase	1.00E-44	SPy1484/ SF370.3
9 (SEQ2080)	2101170	2100967	phage protein [*Streptococcus pyogenes* MGAS9429]	2.00E-13	none	N/A	N/A
10 (SEQ2079)	2100970	2100584	phage protein [*Streptococcus pyogenes* MGAS5005]	5.00E-49	none	N/A	N/A
11 (SEQ2077)	2100440	2100237	phage protein [*Streptococcus pyogenes* MGAS5005]	4.00E-26	none	N/A	N/A
12 (SEQ2076)	2100149	2099850	phage protein [*Streptococcus pyogenes* MGAS5005]	5.00E-25	none	N/A	N/A
13 (SEQ2075)	2099850	2098693	phage protein [*Streptococcus pyogenes* MGAS5005]	0	none	N/A	N/A
14 (SEQ2074)	2098678	2098121	phage protein [*Streptococcus pyogenes* MGAS5005]	2.00E-82	none	N/A	N/A
15 (SEQ2073)	2098078	2096156	putative DNA polymerase A domain - phage associated [*Streptococcus pyogenes* MGAS315]	0	none	N/A	N/A
16 (SEQ2072)	2096151	2093767	putative DNA primase/helicase - phage associated [*Streptococcus pyogenes* MGAS315]	0	none	N/A	N/A

17 (SEQ2071)	2093380	2093105	phage protein, putative [*Alkaliphilus metalliredigenes* QYMF]	1.00E-09	none	N/A	N/A
18 (SEQ2070)	2093108	2091786	phage-related DNA helicase [*Streptococcus pyogenes* MGAS5005]	0	none	N/A	N/A
19 (SEQ2069)	2091785	2091597	phage protein [*Streptococcus pyogenes* MGAS5005]	6.00E-10	none	N/A	N/A
20 (SEQ2067)	2091364	2091179	unknown phage protein [*Streptococcus pyogenes* MGAS10394]	1.00E-21	none	N/A	N/A
21 (SEQ2066)	2091182	2090889	phage protein [Bacteriophage 9429.1]	7.00E-07	none	N/A	N/A
22 (SEQ2065)	2090892	2090176	phage protein [*Streptococcus pyogenes* MGAS6180]	4.00E-23	hypothetical protein - phage associated	4.00E-16	SPy0965/ SF370.2
23 (SEQ2064)	2090179	20899763	phage transcriptional activator [*Streptococcus pyogenes* MGAS5005]	2.00E-65	none	N/A	N/A
24 (SEQ2063)	2089673	2089221	phage Terminase Small Subunit [*Streptococcus pyogenes* MGAS10394]	3.00E-62	putative terminase, small subunit - phage associated	3.00E-07	SPy0971/ SF370.2
25 (SEQ2062)	2089231	2087954	putative terminase large subunit - phage associated [*Streptococcus pyogenes*	0	putative terminase, large subunit phage	8.00E-49	SPy0972/ SF370.2

			MGAS315]		associated		
26 (SEQ2061)	2087938	2086406	phage portal protein, SPP1 family [*Streptococcus agalactiae* H36B]	0	none	N/A	N/A
27 (SEQ2060)	2086446	2084998	phage putative head morphogenesis protein, SPP1 gp7 family domain protein [*Streptococcus agalactiae* H36B]	0	conserved hypothetical protein - phage associated	2.00E-15	SPy0975/ SF370.2
28 (SEQ2059)	2084970	2084782	hypothetical protein SpyM3_1432 [*Streptococcus pyogenes* MGAS315]	1.00E-13	none	N/A	N/A
29 (SEQ2058)	2084777	2084520	hypothetical protein, phage associated [*Streptococcus pyogenes* M1 GAS]	6.00E-42	see Best blastp match	see E value of best match	SPy0685/ SF370.1
30 (SEQ2057)	2084343	2083774	phage scaffold protein [*Streptococcus pyogenes* MGAS5005]	2.00E-62	hypothetical protein - phage associated	4.00E-13	SPy0981/ SF370.2
31 (SEQ2056)	2083761	2082874	phage protein [*Streptococcus pyogenes* MGAS5005]	5.00E-158	putative major head protein, phage associated	0.026	SPy0688/ SF370.1
32 (SEQ2055)	2082862	2082506	putative DNA packaging protein [*Lactobacillus plantarum* bacteriophage phiJL-1]	3.00E-06	conserved hypothetical protein - phage associated	3.00E-05	SPy0986/ SF370.2

33 (SEQ2054)	2082495	2082217	phage protein [*Streptococcus pyogenes* MGAS5005]	2.00E-43	none	N/A	N/A
34 (SEQ2053)	2082220	2081876	phage protein [*Streptococcus pyogenes* MGAS2096]	2.00E-46	conserved hypothetical protein - phage associated	0.039	SPy0988/ SF370.2
35 (SEQ2052)	2081872	2081513	phage protein [*Streptococcus pyogenes* MGAS5005]	4.00E-54	none	N/A	N/A
36 (SEQ2051)	2081501	2080902	major tail protein [*Lactobacillus plantarum* bacteriophage phiJL-1]	1.00E-29	none	N/A	N/A
37 (SEQ2050)	2080848	2080393	phage protein [*Streptococcus pyogenes* MGAS5005]	4.00E-66	none	N/A	N/A
38 (SEQ2049)	2080318	2080085	phage protein [*Streptococcus pyogenes* MGAS5005]	2.00E-23	none	N/A	N/A
39 (SEQ2048)	2080070	2075688	putative tail protein - phage associated [*Streptococcus pyogenes* MGAS315]	0	hypothetical protein, phage associated	1.00E-52	SPy0697/ SF370.1
40 (SEQ2047)	2075676	2074834	minor tail protein [*Lactobacillus plantarum* bacteriophage phiJL-1]	7.00E-11	conserved hypothetical protein - phage associated	0.071	SPy0995/ SF370.2

41 (SEQ2046)	2074824	2072842	phage endopeptidase [*Streptococcus pyogenes* MGAS5005]	0	hypothetical protein, phage associated	1.00E-104	SPy0700/ SF370.1
42 (SEQ2045)	2072842	2071724	hyaluronidase - phage associated [*Streptococcus pyogenes* M1 GAS]	1.00E-149	see Best blastp match	see E value of best match	SPy0997/ SF370.2
43 (SEQ2044)	2071709	2069823	phage infection protein [Bacteriophage 9429.1]	0	hypothetical protein, phage associated	0	SPy0702/ SF370.1
44 (SEQ2043)	2069811	2069374	hypothetical protein - phage associated [*Streptococcus pyogenes* M1 GAS]	4.00E-77	see Best blastp match	see E value of best match	SPy1443/ SF370.3
45 (SEQ2042)	2069377	2068766	hypothetical protein - phage associated [*Streptococcus pyogenes* M1 GAS]	1.00E-94	see Best blastp match	see E value of best match	SPy1001/ SF370.2
46 (SEQ2041)	2068754	2068482	hypothetical protein [*Streptococcus pyogenes* M1 GAS]	4.00E-42	see Best blastp match	see E value of best match	SPy1441/ SF370.1
47 (SEQ2040)	2068485	2068258	putative holin - phage associated [*Streptococcus pyogenes* M1 GAS]	5.00E-27	see Best blastp match	see E value of best match	SPy0707/ SF370.1
48 (SEQ2038)	2068141	2066807	phage-associated lysin [*Streptococcus pyogenes* MGAS6180]	0	putative lysin - phage associated	0	SPy1006/ SF370.2
49 (SEQ2037)	2066635	2065856	enterotoxin [Bacteriophage 9429.2]	4.00E-135	streptococcal exotoxin I	1.00E-129	SPy1007/ SF370.2

50 SEQ2036)	2065830	2065120	streptococcal exotoxin H precursor [*Streptococcus pyogenes* M1 GAS]	1.00E-136	see Best blastp match	see E value of best match	SPy1008/ SF370.2
51 (SEQ2035)	2064262	2064867	putative repressor protein - phage associated [*Streptococcus pyogenes M1 GAS*]	1.00E-04	see Best blastp match	see E value of best match	SPy0939/ SF370.2

Table 3.1: Bioinformatic analysis of the ORFs from the *S. equi* 4047 prophage that carries ORF *seq2045*

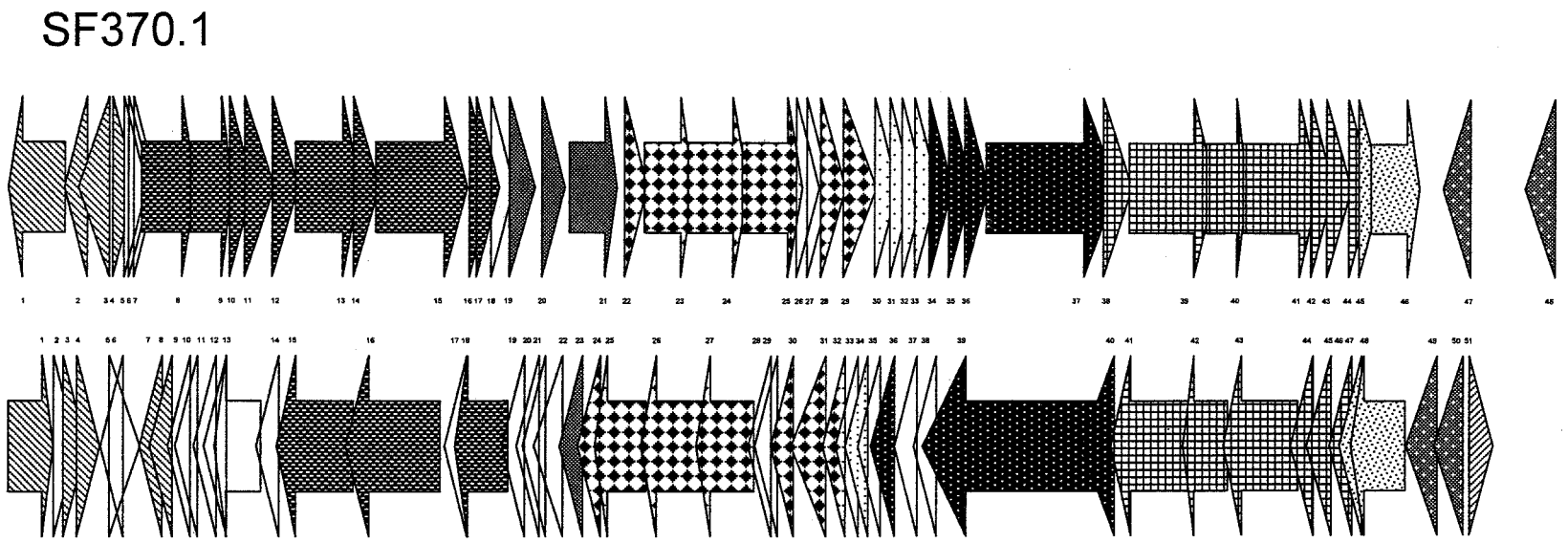

SF370.1
S. equi 4047 prophage that carries ORF seq2045
Lysogeny
Replication
Regulation
Head morphogenesis
Head tail joining
Tail genes
Tail fibres genes
Host lysis
Lysogeny conversion
2 kb

Figure 3.37: The gene maps of the *S. equi* 4047 prophage that carries ORF *seq2045* and the *S. pyogenes* prophage SF370.1. ORFs were not depicted when they were less than 60 codons in length and/or when blastp returned a result of "No significant similarity found". The *S. equi* prophage ORFs were numbered from 1 to 51 (see details in Table 3.1). The *S. pyogenes* ORFs were numbered 1 to 48 (see Canchaya et al.) Genes likely to belong to the same cluster module were marked with the same pattern.

4 Discussion

4.1 HylP1

4.1.1 Quantitative kinetic analysis

Biochemical analysis of HylP1 as a hyaluronate lyase yields a k_{cat} of 7.2 s^{-1} and a K_m of 1.47 mg ml^{-1} or 3.69 mM. In comparison with other unrelated hyaluronate lyases shown in Table 4.1 obvious differences can be seen

Enzyme	Origin	k_{cat} (s^{-1})	*K*m (mg ml^{-1})	*K*m (mM)	pH optimum	Temperature optimum (°C)	Ref
HylP1	Prophage	7.2	1.47	3.69	6.5	37	
TSβH	Prophage	11.1	0.53	1.34	6.5	37	
SEQ2045	Prophage	7.6	0.33	0.83	6.5	37	
H4489A	Prophage	4.9	ND	ND	6	30	(Baker et al., 2002)
hylp2	Prophage	ND	0.31	0.78	6	ND	(Mishra et al., 2006)
S. pneumoniae hyaluronate lyase	Chromosomal	ND	0.04	0.09	6	37	(Nukui et al., 2003)
Serotype III GBS Strain 3502	Chromosomal	8600	ND	ND	6.5	37	(Pritchard et al., 1994)
SpnHL	Chromosomal	ND	0.03	0.08	ND	ND	(Jedrzejas et al., 2002)
SpyHL	chromosomal	15240	ND	ND	6	20	(Jedrzejas et al., 1998)

Table 4.1: Kinetic data of hyaluronan lyases. (ND not determined)

In comparison with other hyaluronate lyases, HylP1 has a relatively low affinity for the substrate requiring approximately 40 times more substrate to reach the half maximum binding. The enzyme is also nearly 2000 times slower in conversion of substrate to product than other unrelated (in terms of sequence similarity) enzymes.
In comparisons such as those made with other hyaluronate lyases it must also be considered that the origin of this enzyme is not bacterial but is in fact bacteriophage. Therefore any comparisons made between it and other bacterial enzymes may not be entirely appropriate. When the enzyme kinetics are compared against other bacteriophage hyaluronate lyases such as the hyaluronate lyase from bacteriophage H4489A of *S. pyogenes* striking similarities can be seen. The k_{cat} of H4489A hyaluronate lyase is 4.9 s^{-1} which is comparable to the k_{cat} of HylP1 at 7.2 s^{-1}. This is indicative of the fact that hyaluronate lyases from bacteriophage sources differ significantly at their catalytic level compared to hyaluronate lyases that are bacterial in origin suggesting a different overall role in degradation processes.

4.1.2 Substrate specificity

Due to similarities between hyaluronan and chondroitin it is often observed that hyaluronate lyases also demonstrate activity against chondroitin. HylP1 however showed no activity against chondroitin even in the presence of excess enzyme. HylP1 therefore suggests specificity for hyaluronan similar to that seen in other bacteriophage encoded hyaluronate lyases HylP2 from *S. pyogenes* bacteriophage 10403 and HylP of *S. pyogenes* bacteriophage H4489A (Baker, 2002, Mishra, 2006) suggesting a different role for this enzyme compared to the less specific hyaluronate lyases that also degrade chondroitin. This substrate specificity has been observed in other bacteriophage encoded hyaluronate lyases.

4.1.3 Mode of action

HPAEC analysis of the products of hyaluronan digestion by HylP1 shows a range of different retention timed peaks which correlate to a range of differently sized oligosaccharides of hyaluronan. This range of product sizes confirms that HylP1 has random endolytic mode of digestion, cleaving at glycosidic bonds at random positions within the polysaccharide chain, producing differently sized oligomers each time.

Of interesting note is the fact that the smallest sized oligomer produced is a tetramer even after exhaustive digestion with excessive enzyme concentrations. A dimmer or trimer is never produced suggesting that a certain degree of binding is needed in the active site for degradation to occur. As a tetramer is the smallest product produced it can be predicted that at least an octamer is needed for complete binding and positioning within the groove which yields two tetramers which can not fully bind for further degradation. A hexamer would not produce the products seen as it would be degraded to yield both a tetramer and a dimer.

Again this can be compared to both other unrelated bacterial hyaluronate lyases and bacteriophage hyaluronate lyases. The hyaluronate lyase from serotype III GBS is a processive mechanism which ultimately yields a dimer as the final product of exhaustive digestion (Pritchard et al., 1994) yet the hyaluronan lyase from *S. pyogenes* bacteriophage H4489A yields predominantly a tetramer like that seen in HylP1.

4.1.4 Site Directed Mutagenesis

Using alignment tool ClustalW, putative key conserved residues within HylP1 were identified as D157 and I156 (Figure 3.1.3). This alignment was performed against a hyaluronidase from bee venom, due to the original belief that HylP1 was also a hyaluronidase. The residues determined by this alignment were conserved and upon study of the 3D structure were shown to fall in the suspected area of the catalytic active site. More commonly implicated in lyases are tyrosine residues and several of these in the same region were also determined and targeted for site directed mutagenesis. The activity of the altered forms of HylP1; D157A, I156A and Y169A,

when compared to native HylP1 were significantly reduced by 88.5, 76.8 and 91.9 % respectively; indicating a possible role in catalysis. Using circular dichroism to ascertain if the loss of activity was due to the loss of catalytic residue or loss of tertiary structure it was apparent that in the case of I156A the lack of activity was due to loss of conformation due to the differences in degree of deflection between it and the native protein under different guanidium chloride concentration. However when the mutants D157A and Y169A were compared against native HylP1 similar curves can be seen suggesting that the conformation is the same between all three forms of the protein which shows that the loss of activity must be due to the loss of catalytic residues.

When these residues are observed within the structure of HylP1 a clear role in the formation of the active site can be seen. The aspartic acid on one strand of the trimer formation lines up with the tyrosine on an adjacent strand at one end of the central triple stranded β helix (Figure 4.1). This forms, on the entire structure, three active sites of a groove topology. The grooves formed between the three strands have every 3-4 residues positively charged arginine or lysine residues which protrude inwards into the groove creating a distance within of approximately 9 – 14 angstroms. This correlates with the size and dimensions of hyaluronan and the positive charge of the groove will readily bind ionically the negative charge of hyaluronan thereby ensuring proper location of the substrate within the active site. The aspartic acid and tyrosine residues within the groove are similar to the pairings seen in other lyases (Lunin et al., 2004) where the tyrosine residue acts as the catalytic base for proton abstraction from the C5 of the uronic acid and the aspartic acid acts to maintain the pKa of the tyrosine. These findings indicate a role for the residues in catalysis of hyaluronan but it must be noted that given the length of the entire groove (~ 65 angstroms) the role of them cannot be excluded as being simply in a binding capacity to ensure proper localisation of the substrate for other residues within the groove.

4.1.5 Triple stranded β helix

The location of the putative catalytic residues within the triple stranded β helix was suggestive of the TSβH being catalytically active. Previously characterised hyaluronate lyases have been shown to be multi domain in nature and that the catalysis of them is located in the α domain. This α domain may only have 10% of the catalytic activity but when combined with the other domains the activity will increase (Akhtar, 2006). Likewise, previously characterised TSβH containing molecules have been catalytic but with the additional presence of accessory globular domains of which none were apparent in HylP1. To ascertain if the catalytic active site of the entire structure did reside within the TSβH a truncated form of the protein was cloned and expressed and still showed full kinetic activity comparable to full length HylP1 (11.1 s^{-1} and 0.53 mg ml^{-1}). This finding indicates a catalytically active TSβH within the bacteriophage encoded HylP1.

4.1.6 Conclusion

From the results discussed here a proposed function of HylP1 can be defined.
The location of HylP1 within the tail fibres of the bacteriophage (Canchaya et al., 2002) is indicative of its role in the invasion of the host due to the tail fibres being the principal region which binds and attacks the host cell membrane. *S. pyogenes* cells are encapsulated in a thick outer layer of hyaluronan and hence the proposed function of HylP1 is in the degradation of this capsule to aid penetration by the bacteriophage of the host cell. The final digestion product of a tetramer suggests that the enzyme serves not the purpose of complete degradation for nutrient utilisation but a role in reducing the viscosity of the capsule allowing the tail spikes to bind to the cell membrane. The relatively poor kinetics of HylP1 support this theory as speed is not essential in this reduction as the bacteriophage can move further through the capsule as conditions become more favourable thereby further anchoring itself into position allowing for more viscosity reduction. The substrate specificity demonstrated by the enzyme against only hyaluronan can be viewed to confirm the theory that the sole purpose of this enzyme is to degrade hyaluronan and is not used as a spreading factor. Unlike other hyaluronate lyases which demonstrate activity

against not only hyaluronan but also against other glycosaminoglycans to facilitate spread of the bacteria, HylP1 shows a degradative activity only against hyaluronan. The groove topology along the TSβH can be assumed to indicate a 'burrowing' mechanism of the tail fibre on which it is found whereby the tip of the fibre will bind and begin digestion and will move through the capsule until a clearer less viscous passage can be made.

The role of bacteriophage in confering virulence on the host bacterium eg Shiga toxin in *E. coli* (Herold et al., 2004) has been previously reported. It could be that HylP1 is utilised by *S. pyogenes* as an additional virulence factor although it could simply remain dormant within *S. pyogenes* and only become active during a bacteriophage infection of *S. pyogenes*. Section 4.2 discusses a HylP1 homologue which has been implicated as a possible virulence factor utilised by its host bacterium. This has implications for HylP1 as it can be proposed that it too behaves as an additional virulence factor.

4.2 SEQ2045

4.2.1 Alignment analysis

A BLAST search of the *S. equi* strain 4047 genome (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_equi) with *hylP1* from the SF370 strain of the M1 *S. pyogenes* serotype revealed one high-scoring segment pair (E probability = 7.6 x 10^{-120}). The *S. equi* sequence of this pair was part of *seq2045*. The proteins encoded by *seq2045* and *hylP1* share 75% amino acid identity. This high identity score indicated that SEQ2045, like HylP1, is likely to be of bacteriophage origin and is probably also comprised of a triple-stranded β-helix, especially as the region of HylP1, which constitutes one polypeptide of the triple-stranded β-helix, amino acids 109-297, is the most conserved stretch of the alignment.

Following successful cloning into expression vector pET-28a the protein was over expressed in *E. coli* BL21 and purified via IMAC allowing characterisation.

4.2.2 Quantitative kinetic analysis

The purified enzyme demonstrated a k_{cat} of of 7.61 s^{-1} and a Km of 0.83 mM. In comparison with HylP1 with a k_{cat} of 7.2 s^{-1} and a Km of 3.69 mM, striking similarities can be seen. As with HylP1, SEQ2045 can be compared against other unrelated hyaluronate lyases and bacteriophage encoded hylauronate lyases. In comparison with bacterial encoded lyases, SEQ2045 is particularly poor with a slow turnover rate and a higher substrate concentration required to reach half saturation suggesting a low affinity. In comparison with bacteriophage encoded hyaluronate lyases such as the hyaluronate lyases from *S. pyogenes* H4489A bacteriophage (k_{cat} of 4.9 s^{-1}) and HylP1 (k_{cat} of 7.2 s^{-1}), SEQ2045 kinetics are similar enough to deem that this enzyme too might be encoded by a bacteriophage integrated within *S. equi*. Other kinetic analysis such as pH optimum, temperature optimum and thermostability were all similar to HylP1, suggesting that SEQ2045 is a bacteriophage encoded hyaluronate lyase. One notably exception between HylP1 and SEQ2045 in terms of kinetic analysis, is the requirement for divalent cations. The

effect of $CaCl_2$ concentration on the rate of activity showed a requirement for approximately 8 mM $CaCl_2$ to reach maximum kinetics. When compared to HylP1 which showed no absolute requirement for additional divalent cations this appears to be a significant difference between the two enzymes. However, purified HylP1 required a buffer containing $CaCl_2$ for stability which SEQ2045 did not. This $CaCl_2$ within the resuspension buffer must have provided HylP1 with sufficient divalent cations for a full rate to be achieved but SEQ2045 which had no $CaCl_2$ in the resuspension buffer would require an additional source within the kinetic reaction.
The role of additional $CaCl_2$ in polysaccharide lyases has been shown to be involved in the neutralization of the carboxyl group of the uronic acid. This is performed by the removal of electrons from the carboxylate oxygen which serves to increase the acidity of the C5 ready for abstraction during the elimination mechanism (Jenkins et al., 2004, Michel et al., 2004).

4.2.3 Mode of action

Studies on the mode of action of SEQ2045 using HPAEC showed similar results as seen with HylP1 with a range of different retention time peaks which correspond to a range of differently sized oligomers. This again shows an endolytic mode of action with the enzyme cleaving at random internal positions. As with HylP1 the smallest sized product of digestion produced was a tetramer unlike other hyaluronate lyases which cause complete degradation to a disaccharide. This result can be interpreted to suggest that SEQ2045 has a similar three dimensional structure to HylP1 with a similar active site topology of a groove in a triple stranded β helix formed by the quarternary structure of the trimerised protein and is involved in the reduction of the viscosity of the HA capsule of streptococci during bacteriophage infection.

SEQ2045 showed no catalytic activity against chondroitin 4-sulphate, chondroitin 6-sulphate and dermatin sulphate similar to that seen in HylP1. The narrow substrate range seen in the bacteriophage-encoded enzymes adds further support to the speculation that prophage-encoded hyaluronate lyases have evolved for the primary purpose of allowing penetration of the bacteriophage through streptococcal HA capsules.

The high degree of similarity between HylP1 and SEQ2045 in regards to sequence, kinetics and mode of action are suggestive of a similar function and structure and hence a similar active site location and composition can be predicted. ClustalW alignment of HylP1 and SEQ2045 showed a good alignment of the aspartic acid and tyrosine residues shown to contribute to the active site of HylP1. These were altered by site directed mutagenesis which ultimately removed the activity of the enzyme with a 91.3 % and 87.6 % loss of activity respectively. The structure of SEQ2045 has yet to be elucidated so the location of these residues within the protein and the putative active site cannot be confirmed but on the basis of similarities between it and HylP1 presumptions can be made suggesting a similar active site topology and role of catalytic residues.

4.2.4 Pathogenic role

Using purified recombinant SEQ2045, the role of the enzyme in pathogenesis during an *S. equi* infection was proposed. Western blot analysis using serum taken from various horses following infection showed that antibodies were raised against the protein during an infection.

Western blots were performed using firstly serum taken from a horse which had not been knowingly been infected with *S. equi*, a horse following a *Rhodococcus equi* infection and a horse following an *S. equi* infection. Using anti horse antibody a strong reaction was seen only in the blot previously incubated in the serum taken from a *S. equi* infected horse. This indicates that the horse infected with *S. equi* had raised antibodies against the protein. Incubation with the serum saw the antibodies bind to the purified protein onto which then bound anti-horse antibodies creating the strong reaction. The lack of a reaction in the serum following a *R. equi* infection shows an extra confirmatory result. If the reaction seen from the *S. equi* infected serum was simply from the horse being hyper immune and the antibodies were directed simply towards anything foreign (in this case the purified protein) then a similar reaction would also be expected in the serum taken from the *R. equi* infected serum, as this too would be in a hyper immune state and antibodies towards any foreign material would be present. This is indicative that the reaction seen against SEQ2045 is unique and targeted specifically towards it.

These first Western blot analyses could however have been flawed in that there is no control in the effect of previous unrecorded infections. All three horses were

different and hence the results seen could be due to differences between the horses. A further western blot analysis using matched pre and post *S. equi* infection serum taken from the same horse removed this element of doubt. Three different matched samples were used and the results showed that in all three horses prior to infection there was no reactivity with SEQ2045 but serum taken following infection showed a high level of reactivity which supports the previous assumption that antibodies are raised against the enzyme during an infection procedure.

4.2.4.1 SEQ2045 as a potential virulence factor

The presence of antibody against SEQ2045 following an *S. equi* infection is suggestive of their expression during pathogenesis. This may be the result of two possible routes.

One may be that the bacteriophage on which the enzyme is suspected to reside is lysogenic and as such may confer its genes and their subsequent products to the host bacterium via horizontal gene transfer. This acquisition of new genes can lead to the evolution of more virulent strains as the new genes may possibly encode proteins which may function as virulence factors. An example of this is seen in *Vibrio cholera*. This bacterium requires cholera toxin to become virulent and the toxin has been shown to be encoded by a bacteriophage (Waldor and Mekalanos, 1996). *S. equi* may be utilising the bacteriophage encoded hyaluronate lyase as a potential virulence factor to use alongside its own hylauronate degrading enzymes for complete degradation. Interestingly, the only non-prophage encoded hyaluronate lyase of *S. equi* strain 4047, SEQ1479, is apparently encoded by a pseudogene that when expressed fails to produce a functioning protein. Therefore, the hyaluronate degrading capabilities of this strain of *S. equi* must come from the prophage encoded genes.

An alternative cause for the results seen in the western blots may be that during an infection, the lysogenic prophage becomes induced and swaps to a lytic lifecycle becoming an active bacteriophage (Broudy et al., 2001). At this point the bacteriophage will be expressing the hyaluronate lyase as it is likely to be involved in capsule degradation allowing for new infection. Therefore it is likely that it will be expressed at high levels during *S. equi* infections and the host may produce not only

antibodies against the bacterium but also against the released bacteriophage and the components of it.

In light of the results it may therefore be necessary to re-evaluate the role of HylP1 from the SF370 strain of *S. pyogenes* serotype M1. It is possible that HylP1, like SEQ2045, is expressed during human *S. pyogenes* infections. HylP1 may therefore also act as an additional virulence factor to HylA, a secreted *S. pyogenes* PL8 enzyme (SPy1032) which has been shown it to be a typical PL8 hyaluronate lyase with activity against HA and chondroitins (unpublished observations). For both *S. pyogenes* and *S. equi*, phage induction *in vivo* may thus lead to the production and release of both phage-encoded virulence factors (e.g. DNAase; superantigens) and phage structural proteins with the capacity to damage connective tissues (i.e. HylP1 and SEQ2045).

Experimental infections of the horse with *S. equi* may therefore be an ideal model for studying the role of hyaluronate lyases in human *S. pyogenes* infections. This therefore provides us with the opportunity to model *S. pyogenes* infections, and in particular the role HA-degrading enzymes play, by performing challenge experiments in horses using *S. equi* mutants lacking SEQ2045. This will ultimately provide vital insights into the influence of hyaluronate lyases in pathogenesis and their possible use as targets for antimicrobial agents and vaccines. Moreover, as the study of *in vivo* interactions between phages and pathogenic bacteria remains an important new frontier for phage biologists, the opportunity to study *S. equi* in the horse may provide an important system for studying the contribution of phages to bacterial virulence in the host.

4.3 Further Work

There are numerous opportunities for future studies of both HylP1 and SEQ2045 as a result of this research. SEQ2045 could be crystallised to allow for the structure to be solved or using the structure of HylP1 as a template, molecular replacement could be carried out to determine the structure of the protein.

Although suspected catalytic residues have been isolated and proven to have an impact on the activity of the enzyme, these residues were determined by alignment with a glycoside hydrolase due to the initial suspected activity of HylP1. Further site directed mutagenesis of HylP1 as a hyaluronate lyase would enhance the understanding of the active site and proposed contribution of the residues. Such residues which could be targeted are asparagine and histidine residues as these have been shown in polysaccharide lyases to contribute to the catalysis. These residues could be determined by studying the three dimensional structure of HylP1 and identifying residues of this nature within the area of the proposed active site. This could also be repeated for SEQ2045 and the subsequent catalytically inactive mutants could be used for the formation of complex crystals with the hyaluronan bound within the active site. With the substrate bound in the active site proposed mechanisms for the catalysis can be determined and the role of the different residues elucidated.

An understanding of the cative site architecture could also give rise to the possibility of designing inhibitors towards this enzyme. The inhibitor designed should be capable of blocking the active site and prevent subsequent catalysis. It should either be non reactive in that it will bind in the active site but is incapable of being degraded or be of such a nature that once degraded it can not be expelled from the active site. One possible suggestion for this may be related to the presence of a negative patch within hyaluronate lyase active sites. This negative patch repels the negative substrate so following catalysis the substrate is expelled from the active site. If the inhibitor could be designed in such a way that it would have a positive charge in the vicinity of this negative patch it would remain within the active site and block subsequent binding of substrate.

SEQ2045 has a clear route of further work. The western blots performed are suggestive that this enzyme is expressed during an *S. equi* infection of horses. A

future study may be to elucidate if this expression is as a result of *S. equi* acquiring this enzyme as an additional virulence factor or if the process of *S. equi* infection causes the prophage to become induced and switches to a lytic lifecycle. If the former is true it poses the distinct possibility of the same being true within *S. pyogenes* which may allow for *S. equi* to be used as a model for *S. pyogenes* infections. Isolation of bacteriophage from S. equi infections would allow for this to be determined. If no phage are present in the infection but the enzyme is clearly expressed it is likely that the host bacterium is using the enzyme for infection purposes. In this instance the possible role of inhibitors becomes more critical as this could well be used as a treatment for infections. The importance of the enzyme during infection could also be determined by using knockout mutations. The gene that encodes the enzyme could be removed from the bacteria and this then used in challenge experiments or in tissue culture experiments to see if it attenuates the pathogenic capabilities of the bacteria in some way. An attenuation of the bacteria following gene knockout would be suggestive of the enzyme playing a pivotal role in the pathogenesis of *S. equi* and in the use of it as a model it would also be true of *S. pyogenes*.

5 References

ACKERMANN, H. W. (2001) Frequency of morphological phage descriptions in the year 2000. Brief review. *Arch Virol,* 146, 843-57.

ACKERMANN, H. W. (2003) Bacteriophage observations and evolution. *Res Microbiol,* 154, 245-51.

ADAMS, J. R., SANDER, G. & BYERS, S. (2006) Expression of hyaluronan synthases and hyaluronidases in the MG63 osteoblast cell line. *Matrix Biol,* 25, 40-6.

AKHTAR, M. S., KRISHNAN, M. Y. & BHAKUNI, V. (2006) Insights into the mechanism of action of hyaluronate lyase. Role of C-terminal domain and Ca^{2+} in the functional regulation of enzyme. *J Biol Chem,* 281, 28336-28344

ALLEN, A. G., LINDSAY, H., SEILLY, D., BOLITHO, S., PETERS, S. E. & MASKELL, D. J. (2004) Identification and characterisation of hyaluronate lyase from Streptococcus suis. *Microb Pathog,* 36, 327-35.

ALLINGHAM, P. G., BROWNLEE, G. R., HARPER, G. S., PHO, M., NILSSON, S. K. & BROWN, T. J. (2006) Gene expression, synthesis and degradation of hyaluronan during differentiation of 3T3-L1 adipocytes. *Archives of Biochemistry and Biophysics,* 452, 83-91.

ANZAI, T., SHEORAN, A. S., KUWAMOTO, Y., KONDO, T., WADA, R., INOUE, T. & TIMONEY, J. F. (1999a) Streptococcus equi but not Streptococcus zooepidemicus produces potent mitogenic responses from equine peripheral blood mononuclear cells. *Vet Immunol Immunopathol,* 67, 235-46.

ANZAI, T., TIMONEY, J. F., KUWAMOTO, Y., FUJITA, Y., WADA, R. & INOUE, T. (1999b) In vivo pathogenicity and resistance to phagocytosis of Streptococcus equi strains with different levels of capsule expression. *Vet Microbiol,* 67, 277-86.

ARESCHOUG, T., CARLSSON, F., STALHAMMAR-CARLEMALM, M. & LINDAHL, G. (2004) Host-pathogen interactions in Streptococcus pyogenes infections, with special reference to puerperal fever and a comment on vaccine development. *Vaccine,* 22 Suppl 1, S9-S14.

ARMING, S., STROBL, B., WECHSELBERGER, C. & KREIL, G. (1997) In vitro mutagenesis of PH-20 hyaluronidase from human sperm. *Eur J Biochem,* 247, 810-4.

BAKER, J. R., DONG, S. & PRITCHARD, D. G. (2002) The hyaluronan lyase of Streptococcus pyogenes bacteriophage H4489A. *Biochem J,* 365, 317-22.

BALI, J. P., COUSSE, H. & NEUZIL, E. (2001) Biochemical basis of the pharmacologic action of chondroitin sulfates on the osteoarticular system. *Seminars in Arthritis and Rheumatism,* 31, 58-68.

BANKS, D. J., BERES, S. B. & MUSSER, J. M. (2002) The fundamental contribution of phages to GAS evolution, genome diversification and strain emergence. *Trends Microbiol,* 10, 515-21.

BAYRAKTAR, H., AKAL, E., SARPER, O. & VARNALI, T. (2004) Modeling glycosaminoglycans - hyaluronan, chondroitin, chondroitin sulfate A, chondroitin sulfate C and keratan sulfate. *Journal of Molecular Structure-Theochem,* 683, 121-132.

BELMIRO, C. L., SOUZA, H. S., ELIA, C. C., CASTELO-BRANCO, M. T., SILVA, F. R., MACHADO, R. L. & PAVAO, M. S. (2005) Biochemical and

immunohistochemical analysis of glycosaminoglycans in inflamed and non-inflamed intestinal mucosa of patients with Crohn's disease. *Int J Colorectal Dis,* 20, 295-304.

BERES, S. B., SYLVA, G. L., BARBIAN, K. D., LEI, B., HOFF, J. S., MAMMARELLA, N. D., LIU, M. Y., SMOOT, J. C., PORCELLA, S. F., PARKINS, L. D., CAMPBELL, D. S., SMITH, T. M., MCCORMICK, J. K., LEUNG, D. Y., SCHLIEVERT, P. M. & MUSSER, J. M. (2002) Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proc Natl Acad Sci U S A,* 99, 10078-83.

BISNO, A. L., BRITO, M. O. & COLLINS, C. M. (2003) Molecular basis of group A streptococcal virulence. *Lancet Infect Dis,* 3, 191-200.

BOSSI, L., FUENTES, J. A., MORA, G. & FIGUEROA-BOSSI, N. (2003) Prophage contribution to bacterial population dynamics. *J Bacteriol,* 185, 6467-71.

BOTZKI, A., RIGDEN, D. J., BRAUN, S., NUKUI, M., SALMEN, S., HOECHSTETTER, J., BERNHARDT, G., DOVE, S., JEDRZEJAS, M. J. & BUSCHAUER, A. (2004) L-Ascorbic acid 6-hexadecanoate, a potent hyaluronidase inhibitor. X-ray structure and molecular modeling of enzyme-inhibitor complexes. *J Biol Chem,* 279, 45990-7.

BOURNE, Y. & HENRISSAT, B. (2001) Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr Opin Struct Biol,* 11, 593-600.

BRADFORD, M. M. (1976) Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry,* 72, 248-254.

BRECHT, M., MAYER, U., SCHLOSSER, E. & PREHM, P. (1986) Increased hyaluronate synthesis is required for fibroblast detachment and mitosis. *Biochem J,* 239, 445-50.

BRICKER, A. L., CAREY, V. J. & WESSELS, M. R. (2005) Role of NADase in virulence in experimental invasive group A streptococcal infection. *Infect Immun,* 73, 6562-6.

BROUDY, T. B., PANCHOLI, V. & FISCHETTI, V. A. (2001) Induction of lysogenic bacteriophage and phage-associated toxin from group a streptococci during coculture with human pharyngeal cells. *Infect Immun,* 69, 1440-3.

BROUDY, T. B., PANCHOLI, V. & FISCHETTI, V. A. (2002) The in vitro interaction of Streptococcus pyogenes with human pharyngeal cells induces a phage-encoded extracellular DNase. *Infect Immun,* 70, 2805-11.

CAMPO, G. M., AVENOSO, A., D'ASCOLA, A., CAMPO, S., FERLAZZO, A. M., SAMA, D. & CALATRONI, A. (2005) Purified human plasma glycosaminoglycans limit oxidative injury induced by iron plus ascorbate in skin fibroblast cultures. *Toxicol In Vitro,* 19, 561-72.

CANCHAYA, C., DESIERE, F., MCSHAN, W. M., FERRETTI, J. J., PARKHILL, J. & BRUSSOW, H. (2002) Genome analysis of an inducible prophage and prophage remnants integrated in the Streptococcus pyogenes strain SF370. *Virology,* 302, 245-58.

CANCHAYA, C., FOURNOUS, G. & BRUSSOW, H. (2004) The impact of prophages on bacterial chromosomes. *Mol Microbiol,* 53, 9-18.

CANCHAYA, C., PROUX, C., FOURNOUS, G., BRUTTIN, A. & BRUSSOW, H. (2003) Prophage genomics. *Microbiol Mol Biol Rev,* 67, 238-76

CASJENS, S. (2003) Prophages and bacterial genomics: what have we learned so far? *Mol Microbiol,* 49, 277-300.

CHANTER, N., SMITH, K. C. & MUMFORD, J. A. (1995) Equine strangles modelled in mice. *Vet Microbiol,* 43, 209-18.

CHANTER, N., TALBOT, N. C., NEWTON, J. R., HEWSON, D. & VERHEYEN, K. (2000) Streptococcus equi with truncated M-proteins isolated from outwardly healthy horses. *Microbiology,* 146 (Pt 6), 1361-9.

CHERR, G. N., YUDIN, A. I. & OVERSTREET, J. W. (2001) The dual functions of GPI-anchored PH-20: hyaluronidase and intracellular signaling. *Matrix Biol,* 20, 515-25.

CHIBANI-CHENNOUFI, S., BRUTTIN, A., DILLMANN, M. L. & BRUSSOW, H. (2004) Phage-host interaction: an ecological perspective. *J Bacteriol,* 186, 3677-86.

CHOCKALINGAM, P. S., ZENG, W., MORRIS, E. A. & FLANNERY, C. R. (2004) Release of hyaluronan and hyaladherins (aggrecan G1 domain and link proteins) from articular cartilage exposed to ADAMTS-4 (aggrecanase 1) or ADAMTS-5 (aggrecanase 2). *Arthritis Rheum,* 50, 2839-48.

COHEN, R., AUJARD, Y., BIDET, P., BOURRILLON, A., BINGEN, E., FOUCAUD, P., FRANCOIS, M., GARNIER, J. M., GENDREL, D., GUILLOT, M., HAU, I., OLIVIER, C., QUINET, B. & RAYMOND, J. (2005) [Streptococcus pyogenes an emerging pathogen]. *Arch Pediatr,* 12, 1065-7.

COHEN, S. N., CHANG, A. C. Y. & HSU, L. (1972) Nonchromosomal Antibiotic Resistance in Bacteria: Genetic Transformation of Escherichia coli by R-Factor DNA.

COURTNEY, H. S., HASTY, D. L. & DALE, J. B. (2006) Anti-phagocytic mechanisms of Streptococcus pyogenes: binding of fibrinogen to M-related protein. *Mol Microbiol,* 59, 936-47.

COUTINHO, P. M. & HENRISSAT, B. (1999) Carbohydrate-Active Enzymes: an integrated database approach. IN GILBERT, H. J., DAVIES, G. J., HENRISSAT, B. & SVENSSON, B. (Eds.) *Recent Advances in Carbohydrate Bioengineering.* Cambridge, The Royal Society of Chemistry.

CSOKA, A. B., FROST, G. I. & STERN, R. (2001) The six hyaluronidase-like genes in the human and mouse genomes. *Matrix Biol,* 20, 499-508.

DATTA, V., MYSKOWSKI, S. M., KWINN, L. A., CHIEM, D. N., VARKI, N., KANSAL, R. G., KOTB, M. & NIZET, V. (2005) Mutational analysis of the group A streptococcal operon encoding streptolysin S and its virulence role in invasive infection. *Mol Microbiol,* 56, 681-95.

DAVIES, G. & HENRISSAT, B. (1995) Structures and Mechanisms of Glycosyl Hydrolases. *Structure,* 3, 853-859.

DAVIES, G. J., MACKENZIE, L., VARROT, A., DAUTER, M., BRZOZOWSKI, A. M., SCHULEIN, M. & WITHERS, S. G. (1998) Snapshots along an enzymatic reaction coordinate: Analysis of a retaining beta-glycoside hydrolase. *Biochemistry,* 37, 11707-11713.

DAY, A. J. & DE LA MOTTE, C. A. (2005) Hyaluronan cross-linking: a protective mechanism in inflammation? *Trends Immunol,* 26, 637-43.

DUDHIA, J. (2005) Aggrecan, aging and assembly in articular cartilage. *Cell Mol Life Sci,* 62, 2241-56.

EFSTRATIOU, A. (2000) Group A streptococci in the 1990s. *J Antimicrob Chemother,* 45 Suppl, 3-12.

ELSAYED, S., HAMMERBERG, O., MASSEY, V. & HUSSAIN, Z. (2003) Streptococcus equi subspecies equi (Lancefield group C) meningitis in a child. *Clin Microbiol Infect,* 9, 869-72.

EVANKO, S. P., ANGELLO, J. C. & WIGHT, T. N. (1999) Formation of hyaluronan- and versican-rich pericellular matrix is required for proliferation and migration of vascular smooth muscle cells. *Arterioscler Thromb Vasc Biol,* 19, 1004-13.

EVANKO, S. P. & WIGHT, T. N. (1999) Intracellular localization of hyaluronan in proliferating cells. *J Histochem Cytochem,* 47, 1331-42.

EYAL, O., JADOUN, J., BITLER, A., SKUTELSKI, E. & SELA, S. (2003) Role of M3 protein in the adherence and internalization of an invasive Streptococcus pyogenes strain by epithelial cells. *FEMS Immunol Med Microbiol,* 38, 205-13.

FERRETTI, J. J., MCSHAN, W. M., AJDIC, D., SAVIC, D. J., SAVIC, G., LYON, K., PRIMEAUX, C., SEZATE, S., SUVOROV, A. N., KENTON, S., LAI, H. S., LIN, S. P., QIAN, Y., JIA, H. G., NAJAR, F. Z., REN, Q., ZHU, H., SONG, L., WHITE, J., YUAN, X., CLIFTON, S. W., ROE, B. A. & MCLAUGHLIN, R. (2001) Complete genome sequence of an M1 strain of Streptococcus pyogenes. *Proc Natl Acad Sci U S A,* 98, 4658-63.

FIEBER, C., PLUG, R., SLEEMAN, J., DALL, P., PONTA, H. & HOFMANN, M. (1999) Characterisation of the murine gene encoding the intracellular hyaluronan receptor IHABP (RHAMM). *Gene,* 226, 41-50.

FLANAGAN, J., COLLIN, N., TIMONEY, J., MITCHELL, T., MUMFORD, J. A. & CHANTER, N. (1998) Characterization of the haemolytic activity of Streptococcus equi. *Microb Pathog,* 24, 211-21.

FLOCK, M., JACOBSSON, K., FRYKBERG, L., HIRST, T. R., FRANKLIN, A., GUSS, B. & FLOCK, J. I. (2004) Recombinant Streptococcus equi proteins protect mice in challenge experiments and induce immune response in horses. *Infect Immun,* 72, 3228-36.

FRANKOVA, J., KUBALA, L., VELEBNY, V., CIZ, M. & LOJEK, A. (2006) The effect of hyaluronan combined with KI(3) complex (Hyiodine wound dressing) on keratinocytes and immune cells. *J Mater Sci Mater Med,* 17, 891-8.

FRASER, J. R., LAURENT, T. C. & LAURENT, U. B. (1997) Hyaluronan: its nature, distribution, functions and turnover. *J Intern Med,* 242, 27-33.

FROST, G. L., CSOKA, T. & STERN, R. (1996) The hyaluronidases: A chemical, biological and clinical overview. *Trends in Glycoscience and Glycotechnology,* 8, 419-434.

GIRISH, K. S., JAGADEESHA, D. K., RAJEEV, K. B. & KEMPARAJU, K. (2002) Snake venom hyaluronidase: an evidence for isoforms and extracellular matrix degradation. *Mol Cell Biochem,* 240, 105-10.

GMACHL, M. & KREIL, G. (1993) Bee venom hyaluronidase is homologous to a membrane protein of mammalian sperm. *Proc Natl Acad Sci U S A,* 90, 3569-73.

GREILING, H., STUHLSATZ, H. W., EBERHARD, T. & EBERHARD, A. (1975) Studies on the mechanism of hyaluronate lyase action. *Connect Tissue Res,* 3, 135-9.

HARRINGTON, D. J., SUTCLIFFE, I. C. & CHANTER, N. (2002) The molecular basis of Streptococcus equi infection and disease. *Microbes Infect,* 4, 501-10.

HASCALL, V. C., MAJORS, A. K., DE LA MOTTE, C. A., EVANKO, S. P., WANG, A., DRAZBA, J. A., STRONG, S. A. & WIGHT, T. N. (2004) Intracellular hyaluronan: a new frontier for inflammation? *Biochim Biophys Acta,* 1673, 3-12.

HASHIMOTO, W., KOBAYASHI, E., NANKAI, H., SATO, N., MIYA, T., KAWAI, S. & MURATA, K. (1999a) Unsaturated glucuronyl hydrolase of Bacillus sp. GL1: novel enzyme prerequisite for metabolism of unsaturated oligosaccharides produced by polysaccharide lyases. *Arch Biochem Biophys,* 368, 367-74.

HASHIMOTO, W., NANKAI, H., SATO, N., KAWAI, S. & MURATA, K. (1999b) Characterization of alpha-L-rhamnosidase of Bacillus sp. GL1 responsible for the complete depolymerization of gellan. *Arch Biochem Biophys,* 368, 56-60.

HELDIN, P. (2003) Importance of hyaluronan biosynthesis and degradation in cell differentiation and tumor formation. *Braz J Med Biol Res,* 36, 967-73.

HENRISSAT, B. (1991) A Classification of Glycosyl Hydrolases Based on Amino-Acid-Sequence Similarities. *Biochemical Journal,* 280, 309-316.

HENRISSAT, B. & DAVIES, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Current Opinion in Structural Biology,* 7, 637-644.

HEROLD, S., KARCH, H. & SCHMIDT, H. (2004) Shiga toxin-encoding bacteriophages - genomes in motion. *International Journal of Medical Microbiology,* 294, 115-121.

HOTEZ, P. J., NARASIMHAN, S., HAGGERTY, J., MILSTONE, L., BHOPALE, V., SCHAD, G. A. & RICHARDS, F. F. (1992) Hyaluronidase from Infective Ancylostoma Hookworm Larvae and Its Possible Function as a Virulence Factor in Tissue Invasion and in Cutaneous Larva Migrans. *Infection and Immunity,* 60, 1018-1023.

HUBER, M., TRATTNIG, S. & LINTNER, F. (2000) Anatomy, biochemistry, and physiology of articular cartilage. *Invest Radiol,* 35, 573-80.

HYNES, W. L., DIXON, A. R., WALTON, S. L. & ARIDGIDES, L. J. (2000) The extracellular hyaluronidase gene (hylA) of Streptococcus pyogenes. *FEMS Microbiol Lett,* 184, 109-12.

HYNES, W. L. & WALTON, S. L. (2000) Hyaluronidases of Gram-positive bacteria. *FEMS Microbiol Lett,* 183, 201-7.

ITOH, T., AKAO, S., HASHIMOTO, W., MIKAMI, B. & MURATA, K. (2004) Crystal structure of unsaturated glucuronyl hydrolase, responsible for the degradation of glycosaminoglycan, from Bacillus sp. GL1 at 1.8 A resolution. *J Biol Chem,* 279, 31804-12.

ITOH, T., HASHIMOTO, W., MIKAMI, B. & MURATA, K. (2006a) Substrate recognition by unsaturated glucuronyl hydrolase from Bacillus sp. GL1. *Biochem Biophys Res Commun,* 344, 253-62.

ITOH, T., OCHIAI, A., MIKAMI, B., HASHIMOTO, W. & MURATA, K. (2006b) A novel glycoside hydrolase family 105: the structure of family 105 unsaturated rhamnogalacturonyl hydrolase complexed with a disaccharide in comparison with family 88 enzyme complexed with the disaccharide. *J Mol Biol,* 360, 573-85.

ITOH, T., OCHIAI, A., MIKAMI, B., HASHIMOTO, W. & MURATA, K. (2006c) Structure of unsaturated rhamnogalacturonyl hydrolase complexed with substrate. *Biochem Biophys Res Commun,* 347, 1021-9.

JEDRZEJAS, M. J., MELLO, L. V., DE GROOT, B. L. & LI, S. (2002) Mechanism of hyaluronan degradation by Streptococcus pneumoniae hyaluronate lyase. Structures of complexes with the substrate. *J Biol Chem,* 277, 28287-97.

JEDRZEJAS, M. J., MEWBOURNE, R. B., CHANTALAT, L. & MCPHERSON, D. T. (1998) Expression and purification of Streptococcus pneumoniae hyaluronate lyase from Escherichia coli. *Protein Expr Purif,* 13, 83-9.

JENKINS, J., SHEVCHIK, V. E., HUGOUVIEUX-COTTE-PATTAT, N. & PICKERSGILL, R. W. (2004) The crystal structure of pectate lyase Pel9A from Erwinia chrysanthemi. *Journal of Biological Chemistry,* 279, 9139-9145.

JENTSCH, H., POMOWSKI, R., KUNDT, G. & GOCKE, R. (2003) Treatment of gingivitis with hyaluronan. *J Clin Periodontol,* 30, 159-64.

KANAMARU, S., LEIMAN, P. G., KOSTYUCHENKO, V. A., CHIPMAN, P. R., MESYANZHINOV, V. V., ARISAKA, F. & ROSSMANN, M. G. (2002) Structure of the cell-puncturing device of bacteriophage T4. *Nature,* 415, 553-557.

KELLY, C., BUGG, M., ROBINSON, C., MITCHELL, Z., DAVIS-POYNTER, N., NEWTON, J. R., JOLLEY, K. A., MAIDEN, M. C. & WALLER, A. S. (2006) Sequence variation of the SeM gene of Streptococcus equi allows discrimination of the source of strangles outbreaks. *J Clin Microbiol,* 44, 480-6.

KUETTNER, K. E. (1992) Biochemistry of articular cartilage in health and disease. *Clin Biochem,* 25, 155-63.

LADLOW, J., SCASE, T. & WALLER, A. (2006) Canine strangles case reveals a new host susceptible to infection with Streptococcus equi. *J Clin Microbiol,* 44, 2664-5.

LAEMMLI, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227, 680-5.

LANNERGARD, J., FLOCK, M., JOHANSSON, S., FLOCK, J. I. & GUSS, B. (2005) Studies of fibronectin-binding proteins of Streptococcus equi. *Infect Immun,* 73, 7243-51.

LAUDER, R. M., HUCKERBY, T. N. & NIEDUSZYNSKI, I. A. (2000) A fingerprinting method for chondroitin/dermatan sulfate and hyaluronan oligosaccharides. *Glycobiology,* 10, 393-401.

LEE, J. Y. & SPICER, A. P. (2000) Hyaluronan: a multifunctional, megaDalton, stealth molecule. *Curr Opin Cell Biol,* 12, 581-6.

LEPPERDINGER, G., MULLEGGER, J. & KREIL, G. (2001) Hyal2--less active, but more versatile? *Matrix Biol,* 20, 509-14.

LI, S. & JEDRZEJAS, M. J. (2001) Hyaluronan binding and degradation by Streptococcus agalactiae hyaluronate lyase. *J Biol Chem,* 276, 41407-16.

LI, S., KELLY, S. J., LAMANI, E., FERRARONI, M. & JEDRZEJAS, M. J. (2000) Structural basis of hyaluronan degradation by Streptococcus pneumoniae hyaluronate lyase. *Embo J,* 19, 1228-40.

LIDEN, A., KARLSTROM, A., LANNERGARD, J., KALAMAJSKI, S., GUSS, B., RUBIN, K. & RYDEN, C. (2006) A fibronectin-binding protein from Streptococcus equi binds collagen and modulates cell-mediated collagen gel contraction. *Biochem Biophys Res Commun,* 340, 604-10.

LIN, B., AVERETT, W. F. & PRITCHARD, D. G. (1997) Identification of a histidine residue essential for enzymatic activity of group B streptococcal hyaluronate lyase. *Biochem Biophys Res Commun,* 231, 379-82.

LIN, G. & STERN, R. (2001) Plasma hyaluronidase (Hyal-1) promotes tumor cell cycling. *Cancer Lett,* 163, 95-101.

LINDMARK, H., NILSSON, M. & GUSS, B. (2001) Comparison of the fibronectin-binding protein FNE from Streptococcus equi subspecies equi with FNZ from S. equi subspecies zooepidemicus reveals a major and conserved difference. *Infect Immun,* 69, 3159-63.

LINKER, A., HOFFMAN, P., MEYER, K., SAMPSON, P. & KORN, E. D. (1960) The formation of unsaturated disacharides from mucopoly-saccharides and their cleavage to alpha-keto acid by bacterial enzymes. *J Biol Chem,* 235, 3061-5.

LOKESHWAR, V. B., YOUNG, M. J., GOUDARZI, G., IIDA, N., YUDIN, A. I., CHERR, G. N. & SELZER, M. G. (1999) Identification of bladder tumor-derived hyaluronidase: its similarity to HYAL1. *Cancer Res,* 59, 4464-70.

LUNIN, V. V., LI, Y. G., LINHARDT, R. J., MIYAZONO, H., KYOGASHIMA, M., KANEKO, T., BELL, A. W. & CYGLER, M. (2004) High-resolution crystal structure of Arthrobacter aurescens chondroitin AC lyase: An enzyme-substrate complex defines the catalytic mechanism. *Journal of Molecular Biology,* 337, 367-386.

MAHONEY, D. J., APLIN, R. T., CALABRO, A., HASCALL, V. C. & DAY, A. J. (2001) Novel methods for the preparation and characterization of hyaluronan oligosaccharides of defined length. *Glycobiology,* 11, 1025-33.

MCDONOUGH, M. A., KADIRVELRAJ, R., HARRIS, P., POULSEN, J. C. & LARSEN, S. (2004) Rhamnogalacturonan lyase reveals a unique three-domain modular structure for polysaccharide lyase family 4. *FEBS Lett,* 565, 188-94.

MEDINA, E. (2004) Models of group A streptococcal diseases: a review of current status. *Drug Discovery Today: Disease Models,* 1, 65-71.

MELLO, L. V., DE GROOT, B. L., LI, S. & JEDRZEJAS, M. J. (2002) Structure and flexibility of Streptococcus agalactiae hyaluronate lyase complex with its substrate. Insights into the mechanism of processive degradation of hyaluronan. *J Biol Chem,* 277, 36678-88.

MICHEL, G., POJASEK, K., LI, Y. G., SULEA, T., LINHARDT, R. J., RAMAN, R., PRABHAKAR, V., SASISEKHARAN, R. & CYGLER, M. (2004) The structure of chondroitin B lyase complexed with glycosaminoglycan oligosaccharides unravels a calcium-dependent catalytic machinery. *Journal of Biological Chemistry,* 279, 32882-32896.

MILLER, J. D. & NEELY, M. N. (2004) Zebrafish as a model host for streptococcal pathogenesis. *Acta Trop,* 91, 53-68.

MISHRA, P., AKHTAR, MS. & BHAKUNI, V. (2006) Unusual structural features of the bacteriophage-associated hyaluronate lyase (hylp2). *J Biol Chem,* 281, 7143-7150

MIYOSHI-AKIYAMA, T., TAKAMATSU, D., KOYANAGI, M., ZHAO, J., IMANISHI, K. & UCHIYAMA, T. (2005) Cytocidal effect of Streptococcus pyogenes on mouse neutrophils in vivo and the critical role of streptolysin S. *J Infect Dis,* 192, 107-16.

MORI, S., AKAO, S., NANKAI, H., HASHIMOTO, W., MIKAMI, B. & MURATA, K. (2003) A novel member of glycoside hydrolase family 88: overexpression, purification, and characterization of unsaturated beta-glucuronyl hydrolase of Bacillus sp. GL1. *Protein Expr Purif,* 29, 77-84.

MUMMERT, M. E. (2005) Immunologic roles of hyaluronan. *Immunol Res,* 31, 189-206.

NAKAI, H., OKUYAMA, M., KIM, Y. M., SABURI, W., WONGCHAWALIT, J., MORI, H., CHIBA, S. & KIMURA, A. (2005) Molecular analysis of alpha-glucosidase belonging to GH-family 31. *Biologia,* 60, 131-135.

NAKAMURA, T., HASEGAWA, T., TORII, K., HASEGAWA, Y., SHIMOKATA, K. & OHTA, M. (2004) Two-dimensional gel electrophoresis analysis of the abundance of virulent exoproteins of group A streptococcus caused by environmental changes. *Arch Microbiol,* 181, 74-81.

NANKAI, H., HASHIMOTO, W., MIKI, H., KAWAI, S. & MURATA, K. (1999) Microbial system for polysaccharide depolymerization: enzymatic route for xanthan depolymerization by Bacillus sp. strain GL1. *Appl Environ Microbiol,* 65, 2520-6.

NELSON, D., SCHUCH, R., ZHU, S., TSCHERNE, D. M. & FISCHETTI, V. A. (2003) Genomic sequence of C1, the first streptococcal phage. *J Bacteriol,* 185, 3325-32.

NISHIMURA, M., YAN, W., MUKUDAI, Y., NAKAMURA, S., NAKAMASU, K., KAWATA, M., KAWAMOTO, T., NOSHIRO, M., HAMADA, T. & KATO, Y. (1998) Role of chondroitin sulfate-hyaluronan interactions in the viscoelastic properties of extracellular matrices and fluids. *Biochim Biophys Acta,* 1380, 1-9.

NUKUI, M., TAYLOR, K. B., MCPHERSON, D. T., SHIGENAGA, M. K. & JEDRZEJAS, M. J. (2003) The function of hydrophobic residues in the catalytic cleft of Streptococcus pneumoniae hyaluronate lyase. Kinetic characterization of mutant enzyme forms. *J Biol Chem,* 278, 3079-88.

OLENINA, L. V., KUZMINA, T. I., SOBOLEV, B. N., KURAEVA, T. E., KOLESANOVA, E. F. & ARCHAKOV, A. I. (2005) Identification of glycosaminoglycan-binding sites within hepatitis C virus envelope glycoprotein E2*. *J Viral Hepat,* 12, 584-93.

OLOFSSON, S. & BERGSTROM, T. (2005) Glycoconjugate glycans as viral receptors. *Ann Med,* 37, 154-72.

OUSKOVA, G., SPELLERBERG, B. & PREHM, P. (2004) Hyaluronan release from Streptococcus pyogenes: export by an ABC transporter. *Glycobiology,* 14, 931-8.

OZIMEK, L. K., KRALJ, S., KAPER, T., VAN DER MAAREL, M. & DIJKHUIZEN, L. (2006) Single amino acid residue changes in subsite - 1 of inulosucrase from Lactobacillus reuteri 121 strongly influence the size of products synthesized. *Febs Journal,* 273, 4104-4113.

PAIVA, P., VAN DAMME, M. P., TELLBACH, M., JONES, R. L., JOBLING, T. & SALAMONSEN, L. A. (2005) Expression patterns of hyaluronan, hyaluronan synthases and hyaluronidases indicate a role for hyaluronan in the progression of endometrial cancer. *Gynecol Oncol,* 98, 193-202.

PERERA, G. & HAY, R. (2005) A guide to antibiotic resistance in bacterial skin infections. *J Eur Acad Dermatol Venereol,* 19, 531-45.

PONNURAJ, K. & JEDRZEJAS, M. J. (2000) Mechanism of hyaluronan binding and degradation: structure of Streptococcus pneumoniae hyaluronate lyase in complex with hyaluronic acid disaccharide at 1.7 A resolution. *J Mol Biol,* 299, 885-95.

PRITCHARD, D. G., LIN, B., WILLINGHAM, T. R. & BAKER, J. R. (1994) Characterization of the group B streptococcal hyaluronate lyase. *Arch Biochem Biophys,* 315, 431-7.

PURE, E. & CUFF, C. A. (2001) A crucial role for CD44 in inflammation. *Trends Mol Med,* 7, 213-21.

RIGDEN, D. J. & JEDRZEJAS, M. J. (2003a) Genome-based identification of a carbohydrate binding module in Streptococcus pneumoniae hyaluronate lyase. *Proteins,* 52, 203-11.

RIGDEN, D. J. & JEDRZEJAS, M. J. (2003b) Structures of Streptococcus pneumoniae hyaluronate lyase in complex with chondroitin and chondroitin sulfate disaccharides. Insights into specificity and mechanism of action. *J Biol Chem,* 278, 50596-606.

RODRIGUEZ, E. & ROUGHLEY, P. (2006) Link protein can retard the degradation of hyaluronan in proteoglycan aggregates. *Osteoarthritis Cartilage,* 14, 823-9.

RYE, C. S. & WITHERS, S. G. (2000) Glycosidase mechanisms. *Current Opinion in Chemical Biology,* 4, 573-580.

SAMBROOK, J. & RUSSELL, D. (2001) Molecular Cloning: A Laboratory Manual, 3rd ed, New York, Cold Spring Harbour Laboratory Press

SATOH, T., NAGAHATA, M., TERAMOTO, A., HATIMORI, A., ABE, K. & IM, S. S. (2004) The basic research on physiological property of functionalized hyaluronan - I. Effect of hyaluronan and sulfated hyaluronan on cell proliferation of human epidermal keratinocytes. *Polymers for Advanced Technologies,* 15, 329-334.

SAUNDERS, J. R., ALLISON, H., JAMES, C. E., MCCARTHY, A. J. & SHARP, R. (2001) Phage-mediated transfer of virulence genes. *Journal of Chemical Technology and Biotechnology,* 76, 662-666.

SCHILLER, J., FUCHS, B. & ARNOLD, K. (2006) The molecular organization of polymers of cartilage in health and disease. *Current Organic Chemistry,* 10, 1771-1789.

SCHRAGER, H. M., ALBERTI, S., CYWES, C., DOUGHERTY, G. J. & WESSELS, M. R. (1998) Hyaluronic acid capsule modulates M protein-mediated adherence and acts as a ligand for attachment of group A Streptococcus to CD44 on human keratinocytes. *J Clin Invest,* 101, 1708-16.

SCHRAGER, H. M., RHEINWALD, J. G. & WESSELS, M. R. (1996) Hyaluronic acid capsule and the role of streptococcal entry into keratinocytes in invasive skin infection. *J Clin Invest,* 98, 1954-8.

SCHULTZ, K., RASMUSSEN, L. M. & LEDET, T. (2005) Expression levels and functional aspects of the hyaluronan receptor CD44. Effects of insulin, glucose, IGF-I, or growth hormone on human arterial smooth muscle cells. *Metabolism,* 54, 287-95.

SCOTT, J. E., THOMLINSON, A. M. & PREHM, P. (2003) Supramolecular organization in streptococcal pericellular capsules is based on hyaluronan tertiary structures. *Exp Cell Res,* 285, 1-8.

SEYFRIED, N. T., MCVEY, G. F., ALMOND, A., MAHONEY, D. J., DUDHIA, J. & DAY, A. J. (2005) Expression and purification of functionally active hyaluronan-binding domains from human cartilage link protein, aggrecan and versican - Formation of ternary complexes with defined hyaluronan oligosaccharides. *Journal of Biological Chemistry,* 280, 5435-5448.

SHAH, D. S. H. & RUSSELL, R. R. B. (2004) A novel glucan-binding protein with lipase activity from the oral pathogen Streptococcus mutans. *Microbiology-Sgm,* 150, 1947-1956.

SHELBURNE, S. A., 3RD, GRANVILLE, C., TOKUYAMA, M., SITKIEWICZ, I., PATEL, P. & MUSSER, J. M. (2005) Growth characteristics of and virulence factor production by group A Streptococcus during cultivation in human saliva. *Infect Immun,* 73, 4723-31.

SMITH, N. L. (2004) Analysis of a family 6 Carbohydrate-Binding module and three family 69 Hyaluronidases. *Applied Sciences.* Newcastle upon Tyne, University of Northumbria.

STERN, R., ASARI, A. A. & SUGAHARA, K. N. (2006) Hyaluronan fragments: an information-rich system. *Eur J Cell Biol,* 85, 699-715.

STUMMEYER, K., DICKMANNS, A., MUHLENHOFF, M., GERARDY-SCHAHN, R. & FICNER, R. (2005) Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. *Nature Structural & Molecular Biology,* 12, 90-96.

TAWADA, A., MASA, T., OONUKI, Y., WATANABE, A., MATSUZAKI, Y. & ASARI, A. (2002) Large-scale preparation, purification, and characterization of hyaluronan oligosaccharides from 4-mers to 52-mers. *Glycobiology,* 12, 421-6.

TERMEER, C., SLEEMAN, J. P. & SIMON, J. C. (2003) Hyaluronan--magic glue for the regulation of the immune response? *Trends Immunol,* 24, 112-4.

THOMASSEN, E., GIELEN, G., SCHUTZ, M., SCHOEHN, G., ABRAHAMS, J. P., MILLER, S. & VAN RAAIJ, M. J. (2003) The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. *Journal of Molecular Biology,* 331, 361-373.

TIMONEY, J. F. (1988) Protecting against 'strangles': a contemporary view. *Equine Vet J,* 20, 392-4.

TIMONEY, J. F. & MUKHTAR, M. M. (1993) The protective M proteins of the equine group C streptococci. *Vet Microbiol,* 37, 389-95.

TINSLEY, C. R., BILLE, E. & NASSIF, X. (2006) Bacteriophages and pathogenicity: more than just providing a toxin? *Microbes Infect,* 8, 1365-71.

TOOLE, B. P. (2001) Hyaluronan in morphogenesis. *Semin Cell Dev Biol,* 12, 79-87.

TSAO, N., TSAI, W. H., LIN, Y. S., CHUANG, W. J., WANG, C. H. & KUO, C. F. (2006) Streptococcal pyrogenic exotoxin B cleaves properdin and inhibits complement-mediated opsonophagocytosis. *Biochem Biophys Res Commun,* 339, 779-84.

VASELLA, A., DAVIES, G. J. & BOHM, M. (2002) Glycosidase mechanisms. *Current Opinion in Chemical Biology,* 6, 619-629.

VOHRA, H., DEY, N., GUPTA, S., SHARMA, A. K., KUMAR, R., MCMILLAN, D. & GOOD, M. F. (2005) M protein conserved region antibodies opsonise multiple strains of Streptococcus pyogenes with sequence variations in C-repeats. *Res Microbiol,* 156, 575-82.

VYNIOS, D. H., KARAMANOS, N. K. & TSIGANOS, C. P. (2002) Advances in analysis of glycosaminoglycans: its application for the assessment of physiological and pathological states of connective tissues. *J Chromatogr B Analyt Technol Biomed Life Sci,* 781, 21-38.

WAGNER, P. L. & WALDOR, M. K. (2002) Bacteriophage control of bacterial virulence. *Infect Immun,* 70, 3985-93.

WALDOR, M. K. & MEKALANOS, J. J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science,* 272, 1910-1914.

WANG, J. Y. & ROEHRL, M. H. (2002) Glycosaminoglycans are a potential cause of rheumatoid arthritis. *Proc Natl Acad Sci U S A,* 99, 14362-7.

WARREN, R. A. J. (1996) Microbial hydrolysis of polysaccharides. *Annual Review of Microbiology,* 50, 183-212.

WEIN, R. O., MCGARY, C. T., DOERR, T. D., POPAT, S. R., HOWARD, J. L., WEIGEL, J. A. & WEIGEL, P. H. (2006) Hyaluronan and its receptors in mucoepidermoid carcinoma. *Head Neck,* 28, 176-81.

WITHERS, S. G. (2001) Mechanisms of glycosyl transferases and hydrolases. *Carbohydrate Polymers,* 44, 325-337.

ZHANG, M., MCDONALD, F. M., STURROCK, S. S., CHARNOCK, S. J., HUMPHERY-SMITH, I. & BLACK, G. W. (2007) Group A streptococcus cell-associated pathogenic proteins as revealed by growth in hyaluronic acid-enriched media. *Proteomics,* 7, 1379-1390.

ZHENG, Y., DENG, X., ZHAO, Y., ZHANG, H. & MARTIN-DELEON, P. A. (2001) Spam1 (PH-20) mutations and sperm dysfunction in mice with the Rb(6.16) or Rb(6.15) translocation. *Mamm Genome,* 12, 822-9.

Appendix A: Chemicals, media and enzymes

A1: Chemicals

Acros

3,5-Dinitrosalicylic acid

Imidazole

BDH Laboratory Supplies

- Bromophenol Blue
- Sodium Sulphite
- Sodium borohydride

BioRad

- Nitrocellulose membrane

Duchefa

- Ethylenediaminetetraacetic acid disodium salt

Fisher BioReagents

- Acryalmide/Bisacrylamide 37.5:1 40 % solution
- Coomassie Blue R-250
- Methanol

Fisher Chemicals

- Acetic acid, glacial
- Ammonium acetate
- Glycerol
- Ammonium persulphate
- N,N,N',N'-tetra methylethylene diamine (TEMED)
- Potassium dihydrogen phosphate
- Dipotassium hydrogen phosphate
- Ethanol
- Sodium hydroxide
- D-(+)-Glucose
- Sodium acetate
- Propan-2-ol

Invitrogen

- BCIP/NBT substrate kit

Melford

Isopropyl-b-D-Thiogalactopyranoside (IPTG)

Ampicillin

Kanamycin

Tris (hydroxymethyl) aminomethane (Tris Base)

Sodium Dodecyl Sulphate

Glycine

N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulphonic acid] (HEPES)

Sodium chloride

Urea

Premier International Foods

Marvel powdered skimmed milk

Sigma

Magnesium sulphate

Magnesium chloride

Potassium chloride

Xylene cyanol

Sodium hyaluronate

Potassium hyaluronate

Phenol

2 mercaptoethanol

Calcium chloride

Bovine serum albumin

Sodium acetate

Cobalt chloride

Cupric chloride

Mangenese chloride

Ponceau S stain

Alkaline phopsphatase conjugated rabbit anti-horse IgG

Nickel (II) chloride

Barium chloride

Strontium chloride

Zinc chloride

Polyoxyethylenesorbitan monooleate (Tween 80)

Bradfords reagent

A2: Media

Oxoid

Agar (Bacteriological agar N° 1)

NZ amine (casein hydrolysate)

Tryptone

Yeast extract

A3: Enzymes

Enzymes and their reaction buffers are shown in tables A3i and A3ii below:

Novagen:

Enzyme	**Buffer composition 1 x**
KOD (DNA polymerase)	Patented

Table A3i: Enzyme and Reaction buffer (1x)

New England Biolabs

Enzyme	**Buffer composition 1 x**
Nde I	Buffer 4 (20 mM Tris-acetate, 10 mM Magnesium acetate, 50 mM Potassium acetate, 1 mM DTT, pH 7.9)
Bam HI	Buffer 3 (100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9)
Dpn I	Buffer 4 (20 mM Tris-acetate, 10 mM Magnesium acetate, 50 mM Potassium acetate, 1 mM DTT, pH 7.9)
T4 DNA ligase	50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 mg ml^{-1} acetylated BSA
Taq (DNA polymerase)	

Table A3ii: Enzymes and reaction buffers (1x)

A4: Kits

Qiagen: QIAprep® Spin Miniprep Kit
QIAquick® Gel Extraction Kit
Qiagen® Plasmid Maxi Kit

Qiagen® kit buffers	Buffer function / components
P1	**Resuspension.** 50 mM Tris-HCl pH8, 10 mM EDTA, 100 µg/ml RNase A
P2	**Lysis.** 200 mM NaOH, 1 % SDS (w/v)
P3	**Neutralization.** 3 M Potassium Acetate pH 5.
N3	**Neutralization.** 25 – 50 % Guanidinium chloride, 10-25 % acetic acid. (Full components unavailable from manufacturer).
EB	**Elution.** Full components not available from the manufacturer.
PE	**Wash.** Full components not available from the manufacturer.
QG	**Agarose gel dissolver** 50 – 100 % Guanidinium chloride. (Full components not available from the manufacturer).
PB	**Wash.** 25 – 50 % Guanidinium chloride, 25 – 50 % isopropanol. (Full components not available from the manufacturer)
QBT	**Equilibration.** 750 mM NaCl, 50 mM MOPS pH 7, 15 % isopropanol (v/v), 0.15 % Triton® X-100 (v/v)
QC	**Wash.** 1 M NaCl, 50 mM MOPS pH 7, 15 % isopropanol (v/v)
QF	**Elution.** 1.25 M NaCl, 50 mM Tris-HCl pH 8.5, 15 % isopropanol (v/v)

Table A4 : Components of Qiagen kit solutions

Appendix B : Equipment

B1 : Autoclaving

Sterilisation was acheived using a benchtop Prestige® Medical 2100 Classic autoclave at 121 °C, 32 lb/inch^2 pressure for 20 min.

B2: Incubators

For the growth of bacteria in liquid culture a Gallenkamp orbital shaker was used. For the growth of bacteria on solid media a Gallenkamp static incubator was used.

B3: Freeze drier

Following dialysis, substrates were lyophilised using the Christ® Alpha 1-2 freeze-drier.

B4: pH meter

When solutions and media were required at a set pH all adjustments were performed using a Jenway Ion Meter 3340 with calibration buffers at pH 4.0, 7.0 and 9.22

B5: Centrifugation

Centrifugation of volumes $\leq$ 1.5 ml was performed using a small Sigma 1-15 bench top microcentrifuge. Volumes > 1.5 ml were centrifuged in a large Sigma 3K18C refrigerated bench top centrifuge using the rotors and inserts given in table B1.

Rotor	Vessel	Insert	RCF	Application
11133	25 ml plastic universal	Clay universal holder (part No 17049)	2000	Harvesting of cells for preparation of competency
11133	30 kDa cut off protein concentrator	Clay universal holder (part No 17049)	2500	Concentration of protein
11133	200 ml Plastic centrifuge vessel (part No 15202)	None	5500	Harvesting of cells
12158	30 ml plastic centrifuge tube (part No 3119-0030)	None	24000	Separation of soluble cell extract from lysed cells
12131	1.5 ml microcentrifuge tube	None	24000	Isolation of DNA from Qiagen Maxi Kit

Table B1: Centifuge rotors, vessels and applications

B6: Agarose gel kits

Electrophoresis of DNA was performed using an Owl Scientific EASY-CAST electrophoresis system with power supplied via an E-C 570-90 E-C Apparatus Corporation power pack.

B7: SDS-PAGE gel kit
Electrophoresis of proteins was performed using a Bio-Rad Mini-PROTEAN 3 Cell kit with power supplied via an E-C 570-90 E-C Apparatus Corporation power pack.

B8: Gel Documentation
Visualisation of agarose and SDS-PAGE gels was achieved using a Bio-Rad Gel Doc 2000 system and Quantity One™ software. Hard copies of the gel picture were printed via a Mitsubishi Video Copy Processor (Model P91), with Mitsubishi thermal paper (K65HM-$_{CE}$/High density type, 110 mm x 21 m).

B9: Sonication
Cell lysis was performed using a MSE Soniprep 150 ultra-sonication machine

B10: Large scale purification of protein
For purification of protein an AKTA Prime plus FPLC system was used and was controlled by PrimeView Software.

B11: Spectrophotometer
For all spectrophotometric measurements a UV Visible Heλios α Spectronic Unicam spectrophotometer was used with an attached heated waterbath for use during assays. When being used for 232 nm assay the spectrophotometer was controlled by Vision 32 software.

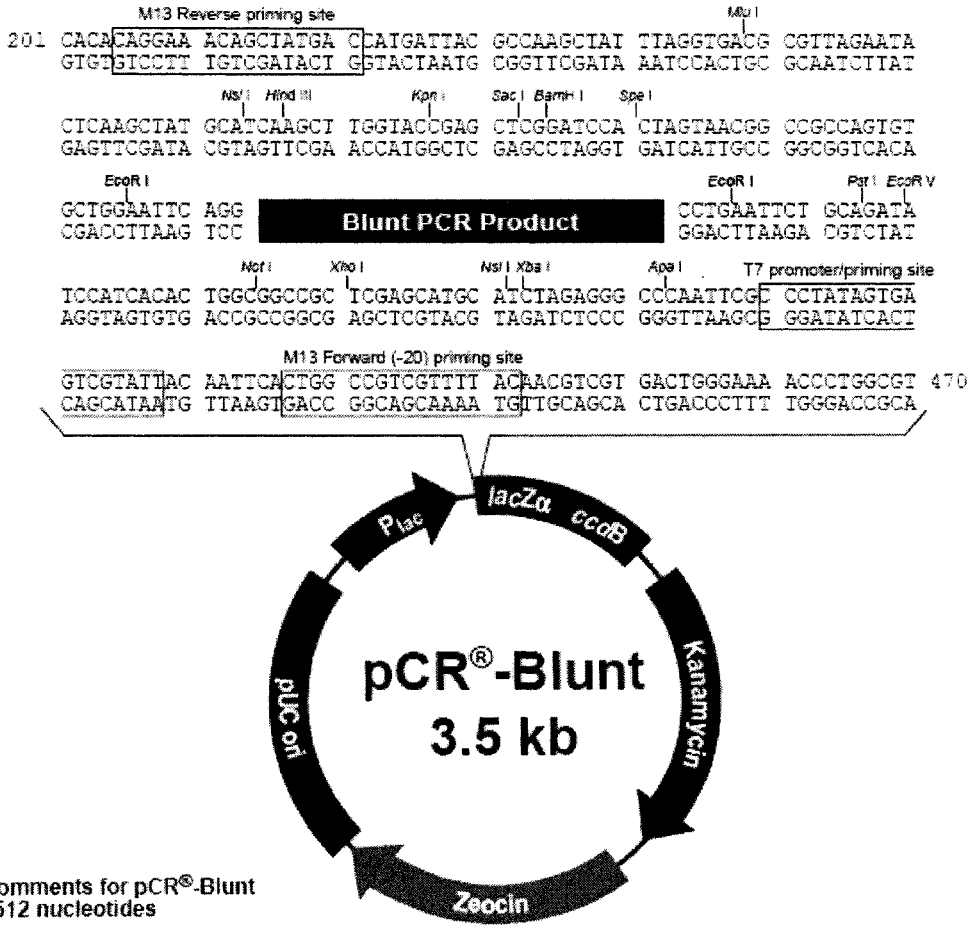
B12: PCR machine
All PCR reactions were performed in an Eppendorf MasterCycler™ PCR machine.

B13: Microtitre plate reader
Measurements were made using a FL600 FAK microtitre plate reader (Bio-Tek Kontron Instruments Inc.) using MicrTek OS software (MTK version 2.5)

B14: Electroporator
Transformation of electrocompetent cells was performed using a Bio-Rad GenePulser Xcell™ with an attached Bio-Rad Pc and CE module and using a Bio-Rad ShcokPod™.

B15: HPAEC
HPAEC was performed using a Dionex system comprised of a LC25 Chromatography Oven, AD25 Absorbance Detector, GP50 Gradient Pump and an AS40 Automated Sampler. The machine was controlled and data gathered using PeakNet software (Dionex)

Appendix C: Vectors



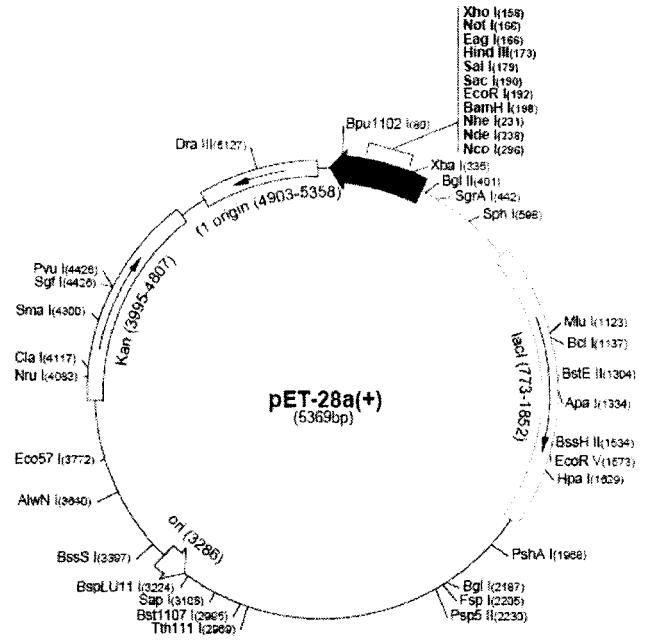
Comments for pCR®-Blunt
3512 nucleotides

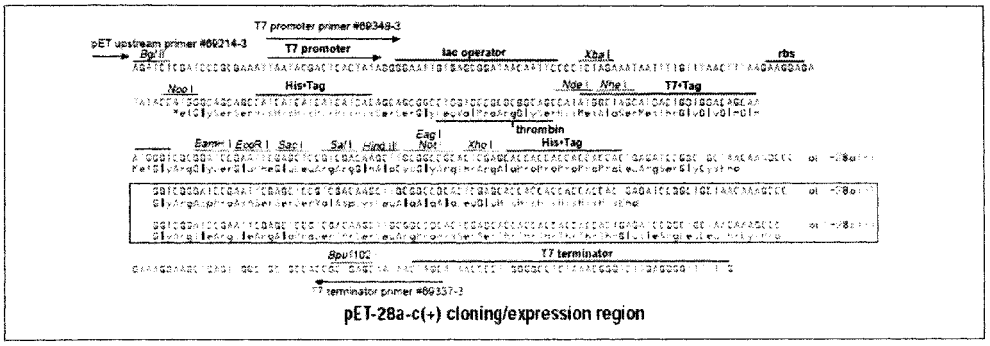
Lac promoter/operator region: bases 95-216
M13 Reverse priming site: bases 205-221
*Lac*Z-alpha ORF: bases 217-570
T7 promoter priming site: bases 400-419
M13 Forward (-20) priming site: bases 427-442
Fusion joint: bases 571-579
*ccd*B lethal gene ORF: bases 580-882
Kanamycin resistance ORF: bases 1231-2025
Zeocin resistance ORF: bases 2231-2605
pUC origin: bases 2673-3386

C1: pCR Blunt vector map

pET-28a(+) sequence landmarks	
T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites	
(*BamH* I - *Xho* I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
f1 origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *BamH* I at 198. pET-28c(+) is a 5367bp plasmid; subtract 2bp from each site beyond *BamH* I at 198.



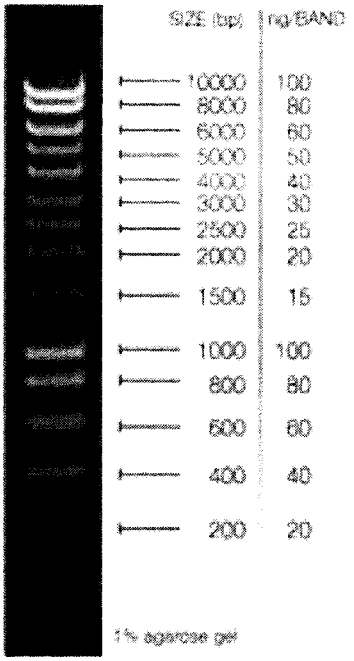


pET-28a-c(+) cloning/expression region

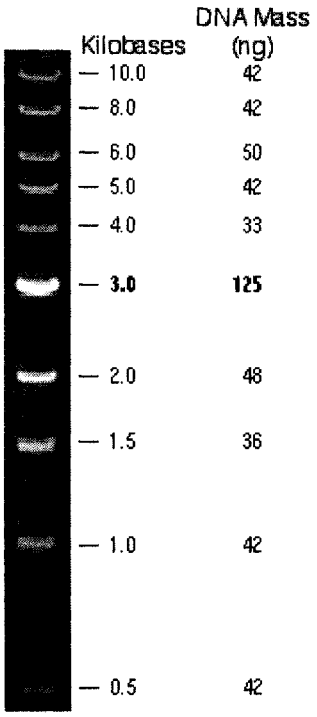
C2: pET-28a vector map

Appendix D: Size standards for electrophoresis

A



B



D1: Agarose gel size standards. A: Bioline Hyperladder I, B: NEB 1 kb ladder

Proteins	Molecular Weight	High Range (S8320)	Wide Range (S8445)	Low Range (M3913)
Myosin from rabbit muscle	200,000	X	X	
β-Galactosidase from *E. coli*	116,000	X	X	
Phosphorylase b from rabbit muscle	97,000	X	X	
Albumin, bovine serum	66,000	X	X	X
Glutamic Dehydrogenase from bovine liver	55,000	X	X	
Ovalbumin from chicken egg	45,000	X	X	X
Glyceraldehyde-3-phosphate Dehydrogenase from rabbit muscle	36,000	X	X	X
Carbonic Anhydrase from bovine erythrocytes	29,000		X	X
Trypsinogen from bovine pancreas	24,000		X	X
Trypsin Inhibitor from soybean	20,000		X	X
α-Lactalbumin from bovine milk	14,200		X	X
Aprotinin from bovine lung	6,500		X	X

D2: SigmaMarkers SDS-PAGE size standards. High and low range molecular weight markers were both used.

Appendix E: Research papers

E1: Smith N.L., Taylor E.J., Lindsay A.-M., Charnock S.J., Turkenburg J.P., Dodson E.J., Davies G.J. & Black G.W., (2005) Structure of a group A streptococcal phage-encoded virulence factor reveals a catalytically active triple stranded β-helix, Proceedings of the National Academy of Science, 102, 17652-17657

E2: Lindsay A.-M., Zhang M., Mitchell Z., Waller A.S., Sutcliffe I.C. & Black G.W. The *Streptococcus equi* prophage-encoded protein SEQ2045 is a hyaluronan-specific hyaluronate lyase that is produced during equine infection. *Submitted to Microbiology*

Structure of a group A streptococcal phage-encoded virulence factor reveals a catalytically active triple-stranded β-helix

Nicola L. Smith*†‡, Edward J. Taylor†§, Anna-Marie Lindsay*, Simon J. Charnock*¶, Johan P. Turkenburg§, Eleanor J. Dodson§, Gideon J. Davies§‖, and Gary W. Black*‖

*Chemical Biology Research Group, School of Applied Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, United Kingdom; and §York Structural Biology Laboratory, Department of Chemistry, University of York, York YO31 5YW, United Kingdom

Edited by Janet Thornton, European Bioinformatics Institute, Cambridge, United Kingdom, and approved October 10, 2005 (received for review June 8, 2005)

Streptococcus pyogenes **(group A Streptococcus) causes severe invasive infections including scarlet fever, pharyngitis (streptococcal sore throat), skin infections, necrotizing fasciitis (flesh-eating disease), septicemia, erysipelas, cellulitis, acute rheumatic fever, and toxic shock. The conversion from nonpathogenic to toxigenic strains of *S. pyogenes* is frequently mediated by bacteriophage infection. One of the key bacteriophage-encoded virulence factors is a putative "hyaluronidase," HylP1, a phage tail-fiber protein responsible for the digestion of the *S. pyogenes* hyaluronan capsule during phage infection. Here we demonstrate that HylP1 is a hyaluronate lyase. The 3D structure, at 1.8-Å resolution, reveals an unusual triple-stranded β-helical structure and provides insight into the structural basis for phage tail assembly and the role of phage tail proteins in virulence. Unlike the triple-stranded β-helix assemblies of the bacteriophage T4 injection machinery and the tailspike endosialidase of the *Escherichia coli* K1 bacteriophage K1F, HylP1 possesses three copies of the active center on the triple-helical fiber itself without the need for an accessory catalytic domain. The triple-stranded β-helix is not simply a structural scaffold, as previously envisaged; it is harnessed to provide a 200-Å-long substrate-binding groove for the optimal reduction in hyaluronan viscosity to aid phage penetration of the capsule.**

crystal | enzyme | lyase | streptococcus | bacteriophage

Streptococcus pyogenes (group A Streptococcus) causes a wide range of acute infections in the respiratory tract and skin with invasive infections causing the life-threatening "flesh-eating" disease, necrotizing fasciitis, and the pyrogenic exotoxin-associated toxic shock syndrome (1). In 1927 it was demonstrated that *S. pyogenes* could be converted from a noninfective to a toxigenic state through the application of a soluble extract of scarlet fever isolates (2), which was subsequently shown to contain bacteriophage. Several of the virulence factors produced by *S. pyogenes* strains are now known to be bacteriophage-encoded. These factors include pyrogenic exotoxins A and C (3, 4), DNase (5), and "hyaluronidase" (6). Indeed, the involvement of phage-encoded virulence factors is now known to be a feature common to many bacterial pathogens (7, 8) with bacteriophage implicated in genome rearrangement (9) and, in the case of *Streptococcus*, interstrain transfer of pyrogenic exotoxin (*Spe*) genes (5, 10, 11). Spes, notably virulence factors SpeC and SpeA, are "superantigens" that directly cross-link the major histocompatibility complex class II on the antigen-presenting cells with T cell receptors, stimulating between 5% and 30% of the total T cell population (5, 12). The increased amounts of inflammatory cytokines result in the symptoms of scarlet fever and, in some cases, toxic shock (1, 12, 13). *Spe* genes are expressed at prophage induction (5), and the simultaneous induction of SpeC and prophage was shown to be mediated by a "soluble factor" by coculturing streptococci with pharyngeal cells (14). The precise events inducing conversion from lysogenic to a lytic cycle remain, however, unclear.

The genome sequencing of S. *pyogenes* strain SF370, revealed three full-length prophage sequences (15) belonging to the *Siphoviridae* family of bacteriophage (λ-like bacteriophage with long noncontractile tails). Of particular interest is the hyaluronidase ORF, HylP1, annotated among the tail fiber genes within the SF370.1 prophage (Fig. 1*a*). The likely role of this hyaluronidase is to allow phage penetration of the streptococcal hyaluronic acid (HA) capsule (16) (Fig. 1*b*) during transfection. To gain insight into the phage-mediated conversion to pathogenicity, in particular the role of the putative tail-fiber hyaluronidase, biochemical and structural analysis of HylP1 (from the only prophage of *S. pyogenes* SF370 that has been shown to be inducible, prophage SF370.1) was undertaken. We confirm that HylP1 is in fact a hyaluronate lyase, catalyzing a β-elimination reaction as first proposed by Baker *et al.* (17) and not a hyaluronan hydrolase. The 3D structure reveals a central triply wound β-helix on a protein >220 Å in length. All previous observations of this unusual fold have been phage-tail proteins, giving considerable credence to a similar role in phage SF370.1. It is widely believed that the triple-stranded β-helix (TSβH) motif is characteristic of structural proteins from viral tails, yet in marked contrast to the triple-stranded β-helical structures of the bacteriophage T4 injection machine (18) and the phages infecting the polysialic-acid-encapsulated human pathogen *Escherichia coli* K1 (19), the catalytic center of HylP1 does not reside on an accessory catalytic domain but on the triple-stranded β-helical core itself, whose organization may shed light on the assembly of the *Siphoviridae* tail fibers themselves.

Experimental Procedures

Cloning, Overexpression, and Site-Directed Mutagenesis. The amino sequence of HylP1 (AAK33657.1) was checked for the presence and location of signal peptide cleavage sites by using the SignalP 3.0 server, (www.cbs.dtu.dk/services/SignalP) (20). The entire coding sequence (CDS) *hylP1* (*SPy0701*) was amplified from the

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: HA, hyaluronic acid; MAD, multiple anomalous difference; TSβH, triple-stranded β-helix.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2C3F).

†N.L.S. and E.J.T. contributed equally to this work.

‡Present address: Institute for Cell and Molecular Biosciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, United Kingdom.

¶Present address: Megazyme International Ireland Limited, Bray, County Wicklow, Republic of Ireland.

‖To whom correspondence may be addressed. E-mail: gary.black@northumbria.ac.uk or davies@ysbl.york.ac.uk.

www.pnas.org/cgi/doi/10.1073/pnas.0504782102

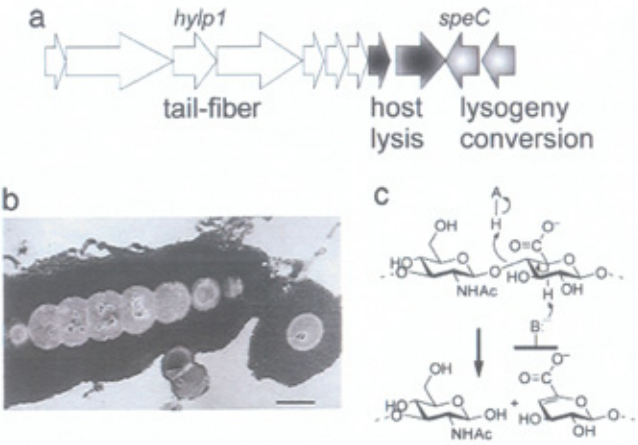


Fig. 1. Partial genetic map of *S. pyogenes* prophage SF370.1 and the hyaluronic acid capsule of streptococci. (*a*) The gene encoding HylP1 is found in the region assigned to the phage tail genes, with *speC*, encoding the phage supertoxin, located in the region associated with lysogenic conversion. (*b*) An electron micrograph of streptococcal cells surrounded by a thick layer of hyaluronan. Typically, the hyaluronan capsule is one to three times the diameter of the cell body. The anionic hyaluronan was visualized with cationic ferritin particles (≈10 nm) containing electron-dense iron cores. (Scale bar: 1 μm, an ≈27,000-fold magnification.) [Reproduced with permission from Glycoforum (www.glycoforum.gr.jp) (Copyright 2004, Glycoforum).] (*c*) The β-elimination reaction catalyzed by a hyaluronate lyase. This reaction results in the formation of a Δ4,5 unsaturated product, which may be detected spectrophotometrically at 232 nm.

genomic DNA of *S. pyogenes* SF370, American Type Culture Collection 700294 (15). This procedure required a two-stage process, because the 3′ end of CDS *hylP3* (*SPy1445*) is identical to *hylP1*. The first set of primers used were 5′-ATggCAATCAATgggTCAAg-3′ and 5′-ggATCCCTATTTTTTTAgTATgAg-3′. The amplified product was cloned into pGEM-T Easy (Promega) and designated pHylP1 + 5′. The *hylp1* ORF was amplified from pHylP1 + 5′ by using 5′-CATATgAgTgAAAATATACCgCTg3-′ and 5′-ggATCCCTATTTTTTTAgTATgAg-3′ and cloned into pET28a (Novagen) on an NdeI–BamHI fragment and designated pHylP1. The conditions used for amplification were as specified for use with Platinum *Pfx* DNA polymerase (Invitrogen). Plasmids pD137A and pY149A were generated from pHylP1 by using the QuikChange site-directed mutagenesis kit (Stratagene) and mutagenic primer pairs 5′-gTggAgCggTCAATATTgCgTTgTCgTCTACCAgAg3-′ and 5′-CTCTggTAgACgACAACgCAATATTgACCgCTCCAC3-′, and 5′-ggTgCTggTgTTgTTgTCgCgTCTgACAATgATACCAgTg-3′ and 5′-CACTggTATCATTgTCAgACgCgACAACAACACCAgCAC C-3′, respectively.

Production and Purification of HylP1 and Derivatives. N-terminally hexahistidine-tagged HylP1 was produced through induction of pHylP1-containing *E. coli* BL21 (DE3; Novagen) cultures grown at 30°C. HylP1 was purified by means of affinity chromatography using nickel-charged Sepharose chelating fast flow resin (GE Healthcare, Chalfont St. Giles, U.K.), followed by gel filtration using a HiLoad 16/60 Superdex 200 prep grade column and concentrated and exchanged into 10 mM Hepes (pH 7.4)/50 mM $CaCl_2$ using a 10-kDa cut-off concentrator unit (Vivascience, Hannover, Germany). D137A and Y149A were produced and purified as for the native enzyme, except they were produced from pD137A- and pY149A-containing *E. coli* BL21, respectively. Seleno-methionine HylP1 was produced and purified as for the native protein, except it was produced from *E. coli* B834 (DE3; Novagen), and 10 mM DTT was incorporated to all buffers before affinity chromatography. Purity of HylP1 was judged by SDS/PAGE, and static nanospray ion trap mass spectrometry (LCQ Advantage, ThermoFinnegan, San Jose, CA) of a tryptic digest was used to confirm the identity of the protein.

HylP1 Assays. Assays were performed spectrophotometrically at 37°C in triplicate at 232 nm by using a 1-cm pathlength quartz cuvette containing 2 mg·ml^{-1} substrate in 50 mM NH_4 Ac buffer (pH 6.5) unless otherwise stated. A molar absorptivity value of 5.5×10^3 liter·mol^{-1}·cm^{-1} was used for concentration calculations (21). All polymeric substrates, hyaluronan (K salt) from umbilical cord, chondroitin 4-sulfate (Na salt) from bovine trachea, and chondroitin 6-sulfate (Na salt) from shark cartilage, were treated with EDTA and dialyzed extensively before use (22). The range of hyaluronan concentrations used for determining k_{cat} and K_m values was 1.3–3 mg·ml^{-1}. Goodness-of-fit statistical analysis of the linear trendlines of the resulting Lineweaver–Burke plots produced R^2 values of ≥0.9935. High-performance anion-exchange chromatography was performed according to the method of Lauder *et al.* (23).

Circular Dichroism (CD). A J-810 spectropolarimeter (Jasco U.K. Ltd., Great Dunmow, Essex, U.K.) was used for CD analysis of native and mutant derivatives of HylP1. CD spectra (260–210 nm) averaged from 15 scans were collected at 20°C by using a bandwidth setting of 3 nm, a scanning speed of 100 nm·min^{-1}, and a 1-mm-pathlength cuvette. Unfolding curves of the proteins in the presence of 0–5 M guanidinium hydrochloride were generated by using the 218-nm ellipticity values (see Fig. 5, which is published as supporting information on the PNAS web site). The protein concentration used was 4.0 mg·ml^{-1} in 10 mM Hepes (pH 7.4)/50 mM $CaCl_2$, and the free energy of the folded and unfolded forms of each protein was determined by least-squares analysis of the data points from the transition region of the unfolding curves of the proteins.

Crystallization and Structure Solution. HylP1 was crystallized from 1 μl of HylP1 (at 10–20 mg·ml^{-1}) with 1–2 μl of the mother liquor (3.25 M sodium formate); 2-methyl-2,4-pentanediol was incorporated into the mother liquor to both improve crystal quality and the ease of freezing. An additional 15% 2-methyl-2,4-pentanediol was included in the stabilizing solution before flash freezing in liquid nitrogen. X-ray data were collected from single crystals at 120 K at the European Synchrotron Radiation Facility. Multiple anomalous difference (MAD) data were collected on beamline ID 14-EH4 and native data on ID 14-EH1. X-ray data were processed with MOSFLM (24), with subsequent computing using programs from the CCP4 suite (25) unless explicitly stated. Crystals belong to the rhombohedral space group H32 with the approximate cell dimensions of $a = b = 58.5$ Å, $c = 586.6$ Å. The structure was solved by using seleno-methionine multiple-wavelength data, with final phasing incorporating native data. Initial phasing used the seleno-methionine MAD data, through MLPHARE, with sites determined by using SHELXD (26). Final phasing incorporated native data, and phases were subsequently improved and extended to 1.8 Å by using DM (27) from the CCP4 suite. Five percent of the data were flagged for cross-validation, and subsequently REFMAC (28) in conjunction with ARPWARP (29) (both CCP4 suite) were used to build the sequence into the electron density automatically. QUANTA and X-FIT (Accelrys, Inc., San Diego) were used to make manual corrections to the model. Solvent molecules were added using X-SOLVATE routines of XFIT and checked manually. The final structure, refined at 1.8-Å resolution, has $R_{cryst} = 0.19$, $R_{free} = 0.23$ with deviations from stereochemical target values of 0.012 Å for bonds and 1.3° for angles (Table 1).

Table 1. X-ray data and structure quality statistics for the *S. pyogenes* SF370 HylP1

Statistic	HylP1 native	HylP1 MAD peak	HylP1 MAD remote	HylP1 MAD inflection
ESRF radiation source	ID 14-EH1	ID 14-EH4	ID 14-EH4	ID 14-EH4
Wavelength, Å	0.93400	0.97950	0.93930	0.97960
Resolution of data (outer shell), Å	195.52–1.80 (1.90–1.80)	43.44–2.50 (2.64–2.50)	50.64–2.50 (2.64–2.50)	196.27–2.50 (2.64–2.50)
R_{merge}*	0.097 (0.259)	0.062 (0.145)	0.064 (0.223)	0.055 (0.153)
Mean $I/\sigma I$ (outer shell)	14.0 (2.8)	18.3 (7.4)	19.4 (5.2)	19.7 (6.9)
Completeness, %	98.8 (98.8)	99.9 (99.9)	99.9 (99.9)	99.8 (99.8)
Multiplicity	5.5 (4.1)	5.2 (5.3)	5.7 (5.7)	5.0 (5.0)
No. protein atoms	2479			
No. ligand atoms	2 (Na^+ ions)			
No. solvent waters	267			
R_{cryst}†	0.19			
R_{free}	0.22			
rms deviation 1–2 bonds, Å	0.012			
Rms deviation 1–3 bonds, °	1.346			
Avg. protein B, Å²	29			
Avg solvent B, Å²	39			

Outer resolution bin statistics are given in parentheses. ESRF, European Synchrotron Radiation Facility.

*$R_{merge} = \Sigma_{hkl}\Sigma_i|I_{hkli} - \langle I_{hkli}\rangle|/\Sigma_{hkl}\Sigma_i\langle I_{hkli}\rangle$.

†$R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$.

Results and Discussion

Carbohydrate-active enzymes, notably glycoside hydrolases, carbohydrate esterases, glycosyltransferases, and polysaccharide lyases, have been grouped together into a large number of sequence-based families. There are almost 100 glycoside hydrolase families but merely 15 families of polysaccharide lyase, this discrepancy reflecting the absolute requirement for uronic-acid substrates for the latter enzyme class. Prior to this work, glycoside hydrolase CAZy (http://afmb.cnrs-mrs.fr/CAZY) family 69 (GH69) contained 26 ORFs, all from bacteriophage that infect streptococcal strains (predominantly *S. pyogenes* but also *S. equi*) that were predominantly annotated as hydrolytic hyaluronidases (EC 3.2.1.35). A feature of this enzyme family is its very high degree of sequence conservation

Phage Infection of *S. pyogenes* Involves the Action of an *Endo*-Acting Hyaluronan-Specific Lyase. *The S. pyogenes* M1 group A Streptococcus SF370 virulence factor hyaluronidase, HylP1, was expressed as a full-length 38.4-kDa N-terminally hexahistidine-tagged protein. HylP1 is active in the degradation of HA with a k_{cat} of 7.2 ± 0.9 s^{-1} and K_m = 1.47 ± 0.02 mg·ml^{-1}. Increase in absorbance at 232 nm is consistent with the action of a lyase (β-eliminase) whose action results in the production of a Δ4,5 unsaturated product (Fig. 1*c*). The enzyme is not a hydrolase, as implied by the enzyme commission number, but instead HylP1 [and its sequence homologs such as the *S. pyogenes* bacteriophage H4489A enzyme (17)] are polysaccharide lyases. Consequently, all the former GH69 enzymes have been reassigned to a new PL family, PL16.

In contrast to the unrelated hyaluronate lyases from group B streptococci, whose function is the degradation of the host-organism connective tissue to facilitate pathogen invasion (31–34), HylP1 is inactive on other connective tissue glycosaminoglycans such as chondroitin 4-sulfate and chondroitin 6-sulfate. The *in vivo* function of HylP1 is further revealed by product analysis. In stark contrast to the group B streptococcal enzymes, indeed to any previously characterized carbohydrate-active enzymes, the end products of HylP1 digestion of HA are long oligosaccharides, predominantly ΔHA_6 and ΔHA_8, with small amounts of ΔHA_4, ΔHA_{10} and longer (Fig. 2). HylP1 thus has an *endo* (random internal cut) mechanism of action with a greatly extended substrate-binding region, which requires at least four to six kinetically significant subsites on either side of the cleavage site. The function of HylP1 is therefore unlikely to be the complete degradation of connective tissue, or even HA itself. Instead, it would appear that the role of HylP1 is the introduction of widely spaced cuts in HA to facilitate a local reduction in capsule viscosity to aid phage invasion of the bacterial host.

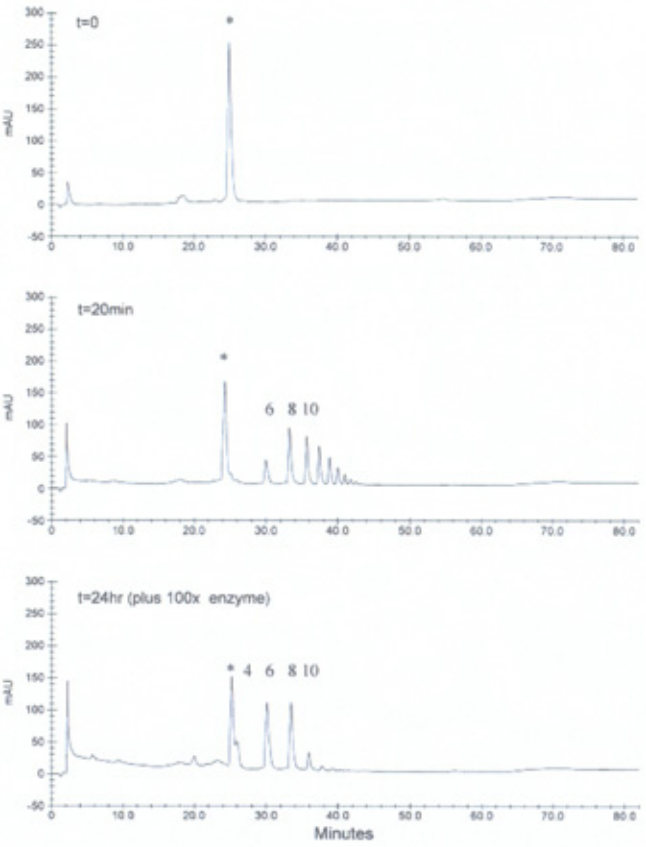


Fig. 2. Catalytic activity of *S. pyogenes* HylP1. High-performance anion exchange chromatography (see ref. 23) analysis of the products of a digestion of hyaluronic acid with HylP1 (20 min incubation with HylP1 at 2 $\mu g \cdot ml^{-1}$). Identification of hyaluronan oligosaccharides, over a range of time points from 0 min to 24 h, was achieved by using electrospray mass spectrometry (42) with the degree of polymerization of product shown. *, artifact.

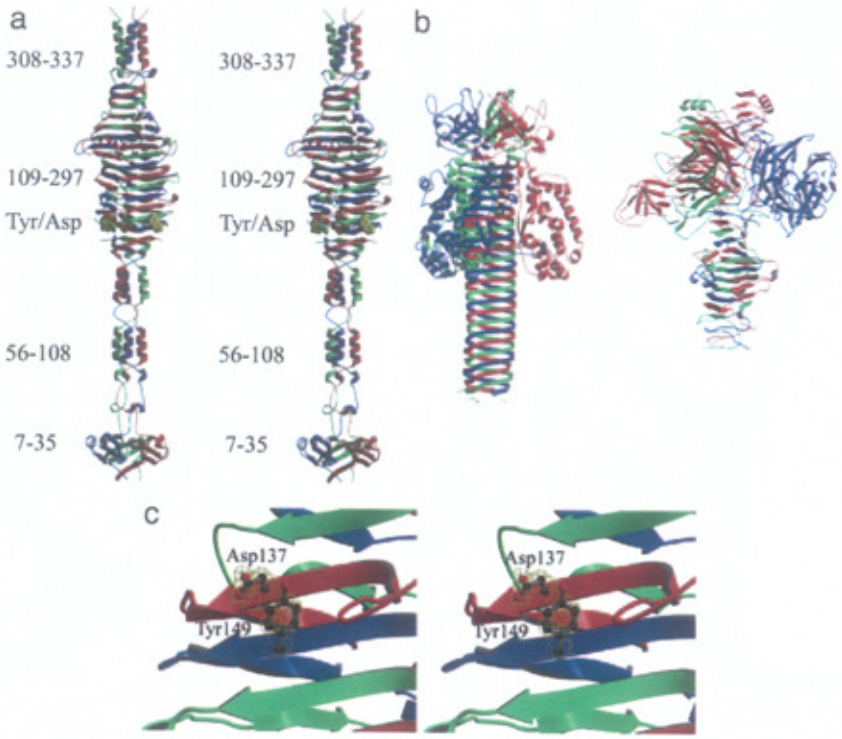


Fig. 3. The 3D structure of HylP1. (*a*) Divergent (wall-eyed) stereoview of HylP1 colored according to chain with the Asp-137/Tyr-149 couple shown in cpk representation (yellow). (*b*) The TSβHs of the bacteriophage T4 gp5 and gp12 proteins and the *E. coli* K1 bacteriophage poly(2,8-sialidase). HylP1 is unique in possessing catalytic activity within the TSβH itself, in marked contrast to the appended catalytic lysozyme and sialidase of the bacteriophage T4 and *E. coli* K1 bacteriophage proteins. (*c*) Close-up of the Asp-137/Tyr-149 pair whose mutation renders HylP1 inactive while still maintaining structural integrity. These images were drawn with MOLSCRIPT (43) and BOBSCRIPT (30).

The 3D Structure of HylP1 Reveals a Catalytically Active TSβH. The 3D structure of HylP1 was solved by using three-wavelength multiple-wavelength anomalous dispersion data from a selenomethionyl-derived protein, and harnessing native data to 1.8 Å, HylP1 crystallizes in the H32 (hexagonal setting) rhombohedral cell with unusual cell dimensions of $a = b = 58.5$ Å, $c = 586.6$ Å. The structure of HylP1 is composed of three intertwined polypeptides centered on a crystallographic threefold axis. The structure (Fig. 3*a*) can be traced from residue 6 through to the C terminus, with the sole exception of the loop from Ala-122 to Ser-128, which is disordered and which may (see below) also be implicated in catalysis. The N-terminal domain (7–35) forms a mixed globular α/β capping region that is followed by a coiled region and segmented α-helical coiled-coils to residue 108. The central "core" of HylP1 then follows, a single extended right-handed TSβH in the form of an irregular triangular tube, where each of the three faces of the tube is composed of alternating β-strands from each of the three polypeptides, with the β-strands orthogonal to the long axis of the enzyme. The faces of the strands are deeply concave, contributing to an extended polysaccharide-

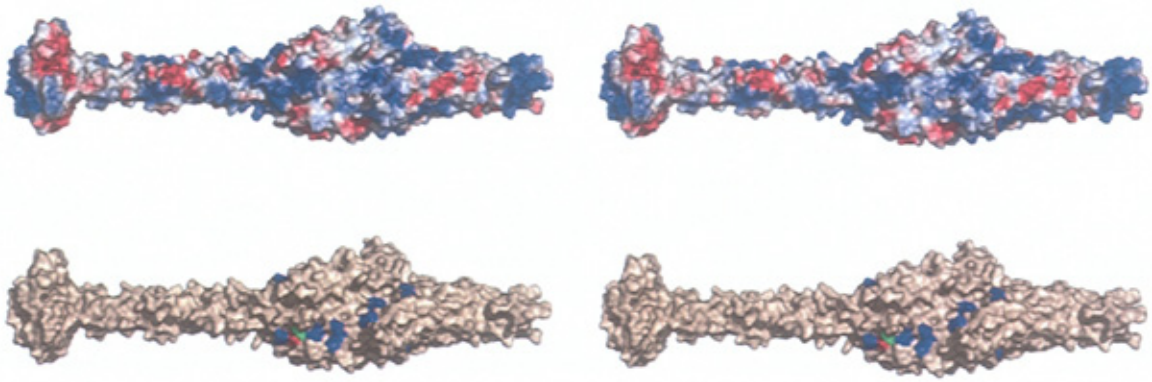


Fig. 4. The substrate-binding groove of HylP1. (*Upper*) Electrostatic surface potential of HylP1 (in divergent stereo) revealing an extended substrate binding cleft that is predominantly positively charged, complementing the negative charge of the substrate hyaluronic acid. (*Lower*) Complementarity with the alternating glucuronic acid moieties of the substrate along the putative substrate groove is provided by a "barber's pole" of positively charged residues (colored blue) including Arg-163, -190, -216, -277, and -279 and Lys-117, -119, -166, -176, and -226. Asp-137 and Tyr-149 are colored red and green, respectively. This figure was drawn with PYMOL (http://pymol.sourceforge.net).

binding surface (Fig. 4; discussed below). The TSβH of HylP1 is composed of residues Lys-108 through to Leu-281 and is ≈65 Å in length. The repeat on each side of the TSβH varies between 8 and 19 residues per turn. At the C terminus of the TSβH there is a short stretch from Ser-282 to Lys-297 where the β-sheets do not intertwine but interdigitate. The remainder of the 220-Å-long structure is composed of random coil and an α-helical "nose."

To our knowledge, the TSβH structural element has been observed in three other bacteriophage tail proteins: the T4 short tail fiber gp12 (35) and cell-puncturing device gp5 (36) and the core structure of the *E. coli* K1 bacteriophage K1F endo-2,8-sialidase (19). The latter two proteins also are involved in phage infection of the bacterial host through degradation of the cell-wall peptidoglycan and capsular polysialic acid, respectively. Threefold symmetry is thus a feature of tail-fiber proteins, manifesting itself in another such protein, tailspike gp9 of bacteriophage P22, except in this case not as a TSβH but as three single-stranded parallel β-helices (37). The width of each face of the triangular tube formed by the HylP1 TSβH varies from 49 Å at the widest point (Asn-274) to 20 Å at its narrowest. Thus, in contrast to gp5 and gp12, whose shape is regular and "smooth," the TSβH domain of HylP1 is both irregular and concave.

Beyond the simple topology, HylP1 is dissimilar to any reported protein structures (all DALI alignments have *Z*-scores of ≤1.2), and all primary sequence alignments of significance (*E*-values < 1.7) were also of phage-encoded hyaluronate lyases of *S. pyogenes* strains. This protein structure is therefore highly unique and highly specialized to *S. pyogenes* bacteriophage. The fact that there is no structural similarity between the TSβH domain of gp5 and gp12 with that of HylP1 and that the bacteriophage that encode these proteins, T4 and SF370.1, respectively, belong to different families suggests that the utilization of this structural element by tail-fiber proteins may be an example of convergent evolution. Indeed, in a topographical, as opposed to topological, sense, HylP1 is more similar to the single-stranded parallel β-helices of known carbohydrate-active enzymes such as several pectinases (for example, ref. 38), chondroitinase B (39), and tailspike gp9 (37). These domains are irregular, contain grooves, are catalytically active, and, in the case of gp9, also a component of a noncontractile phage tail.

Catalytic Center of HylP1. Perhaps the single most important difference between HylP1 and TSβH domains observed previously is that HylP1 catalysis requires no additional catalytic domains, such as the T4 lysozyme domain appended to gp12 and the bacteriophage K1F endosialidase (Fig. 3*b*). We have not obtained complex crystals of HylP1, but we are able to provide insight into the regions critical for catalysis. A triple-helical domain-truncated construct, from residues Asn-106–Asp-310, is catalytically active, confirming that the catalytic center resides in the TSβH domain itself. The TSβH domain provides an extended groove with which to bind long stretches of HA. Every three to four β-strands, positively charged Arg or Lys residues protrude into this grove, spaced between 9 and 14 Å apart, consistent with the spacing of the negatively charged glucuronic acid of the alternating GlcA-GlcNAc polymer that is HA (Fig. 4). Closer inspection of the structure of HylP1 revealed an Asp–Tyr pair (Fig. 3*c*), Asp-137, Tyr-149 (adjacent to a loop that is disordered in the native structure), reminiscent of similar pairings at the center of structurally unrelated hyaluronate and chondroitin lyases in which the Tyr residues act as the catalytic base for proton abstraction from the C5 of the uronic acid with the Asp involved in p*K*a maintenance (for example, ref. 40). Tyr-149 is positioned ≈20 and 60 Å from each end of the TSβH, consistent with the location of the active center determined through product analysis. Both D137A and Y149F mutants are completely inactive, whereas CD analysis indicated that there are no significant structural differences between the native and mutant proteins; indeed, the free energies of the folded and unfolded forms of each protein are also very similar at approximately -1 kcal·mol^{-1}. We therefore propose that Asp-137 and Tyr-149 have a catalytic function, but given the length of the binding groove, we cannot exclude a solely binding role for these residues.

The unique domain-structure of HylP1, viewed in light of the gp12 and gp9 structures, allows speculation on a possible orientation within the phage tail. The N-terminal globular domain of HylP1 resembles the globular collar domain of gp12 (41). The gp12 globular collar domain is adjacent to its receptor-binding domain, located at the C terminus of gp12. HylP1 has no such receptor-binding domain, because the globular domain of HylP1 is at the N terminus of the enzyme. This distinction is important because it does not rule out the possibility that there is another protein linked to HylP1 by means of the N-terminal globular domain, and hence the tail fiber of SF370.1 may not terminate with HylP1.

HylP1 is a previously undescribed catalytically active TSβH protein. The structure confirms its origins as part of a bacteriophage tail assembly whose function is as an "enzymatic needle" that digests the hyaluronic acid capsule during phage invasion of the *S. pyogenes* host. The role of HylP1 is not efficient hyaluronan degradation *per se* but efficient reduction of the viscosity of hyaluronan in the immediate vicinity of a single phage particle, to allow phage penetration to the cell wall. HylP1 has therefore recruited an extremely long structural protein whose extended binding surface provides exactly the unusual properties demanded for phage penetration of the capsule. The concave sides of the TSβH harness an exceedingly large number of subsites, which both prevent the enzyme from becoming "side-tracked" by complete degradation of the hyaluronic acid and help avoid product inhibition. It is particularly significant that the catalytic center resides within the TSβH. It had been assumed, at least tacitly, that the TSβH domain only provided a structural scaffold onto which catalytic domains were subsequently grafted, as illustrated by the bacteriophage T4 injection apparatus and the *E. coli* K1 phage-encoded sialidase. The HylP1 structure demonstrates that this assumption is not necessarily the case and suggests that catalytically active TSβHs will exist elsewhere. Given the increasing resistance to conventional antibiotics, we hope that insight into phage-host specificity and infection processes will have an impact on the subsequent design of novel bacteriophage-derived antibiotics.

We thank Ms. Meng Zhang for technical help with the mass spectrometry, and the European Synchrotron Radiation Facility for provision of beamlines. G.J.D. is a Royal Society University Research Fellow.

1. Cunningham, M. W. (2000) *Clin. Microbiol. Rev.* **13,** 470–511.
2. Frobisher, M. & Brown, J. (1927) *Bull. Johns Hopkins Hospital* **41,** 167–173.
3. Bohach, G. A., Fast, D. J., Nelson, R. D. & Schlievert, P. M. (1990) *Crit. Rev. Microbiol.* **17,** 251–272.
4. Weeks, C. R. & Ferretti, J. J. (1984) *Infect. Immun.* **46,** 531–536.
5. Broudy, T. B., Pancholi, V. & Fischetti, V. A. (2002) *Infect. Immun.* **70,** 2805–2811.
6. Hynes, W. L. & Ferretti, J. J. (1989) *Infect. Immun.* **57,** 533–539.
7. Canchaya, C., Fournous, G., Chibani-Chennoufi, S., Dillmann, M. L. & Brussow, H. (2003) *Curr. Opin. Microbiol.* **6,** 417–424.
8. Casjens, S. (2003) *Mol. Microbiol.* **49,** 277–300.
9. Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y., Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T., *et al.* (2003) *Genome Res.* **13,** 1042–1055.
10. Broudy, T. B. & Fischetti, V. A. (2003) *Infect. Immun.* **71,** 3782–3786.
11. Cleary, P. P., LaPenta, D., Vessela, R., Lam, H. & Cue, D. (1998) *Infect. Immun.* **66,** 5592–5597.
12. Papageorgiou, A. C. & Acharya, K. R. (2000) *Trends Microbiol.* **8,** 369–375.
13. Marrack, P. & Kappler, J. (1990) *Science* **248,** 705–711.
14. Broudy, T. B., Pancholi, V. & Fischetti, V. A. (2001) *Infect. Immun.* **69,** 1440–1443.

15. Ferretti, J. J., McShan, W. M., Ajdic, D., Savic, D. J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, A. N., Kenton, S., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98,** 4658–4663.
16. Hynes, W. L., Hancock, L. & Ferretti, J. J. (1995) *Infect. Immun.* **63,** 3015–3020.
17. Baker, J. R., Dong, S. & Pritchard, D. G. (2002) *Biochem. J.* **365,** 317–322.
18. Rossmann, M. G., Mesyanzhinov, V. V., Arisaka, F. & Leiman, P. G. (2004) *Curr. Opin. Struct. Biol.* **14,** 171–180.
19. Stummeyer, K., Dickmanns, A., Muhlenhoff, M., Gerardy-Schahn, R. & Ficner, R. (2005) *Nat. Struct. Mol. Biol.* **12,** 90–96.
20. Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. (2004) *J. Mol. Biol.* **340,** 783–795.
21. Yamagata, T., Saito, H., Habuchi, O. & Suzuki, S. (1968) *J. Biol. Chem.* **243,** 1523–1535.
22. Brown, I. E., Mallen, M. H., Charnock, S. J., Davies, G. J. & Black, G. W. (2001) *Biochem. J.* **355,** 155–165.
23. Lauder, R. M., Huckerby, T. N. & Nieduszynski, I. A. (2000) *Glycobiology* **10,** 393–401.
24. Leslie, A. G. W. (1992) in *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, eds. Wolf, W. M. and Wilson, K. S. (Daresbury Lab., Warrington, U.K.), Vol. 26.
25. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D* **50,** 760–763.
26. Schneider, T. R. & Sheldrick, G. M. (2002) *Acta Crystallogr. D* **58,** 1772–1779.
27. Cowtan, K. D. & Main, P. (1996) *Acta Crystallogr. D* **49,** 148–157.
28. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) *Acta Crystallogr. D* **53,** 240–255.
29. Perrakis, A., Morris, R. & Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463.
30. Esnouf, R. M. (1997) *J. Mol. Graphics. Model* **15,** 132–134.
31. Baker, J. R. & Pritchard, D. G. (2000) *Biochem. J.* **348,** 465–471.
32. Shain, H., Homer, K. A. & Beighton, D. (1996) *J. Med. Microbiol.* **44,** 372–380.
33. Pritchard, D. G., Lin, B., Willingham, T. R. & Baker, J. R. (1994) *Arch. Biochem. Biophys.* **315,** 431–437.
34. Nukui, M., Taylor, K. B., McPherson, D. T., Shigenaga, M. K. & Jedrzejas, M. J. (2003) *J. Biol. Chem.* **278,** 3079–3088.
35. van Raaij, M. J., Schoehn, G., Burda, M. R. & Miller, S. (2001) *J. Mol. Biol.* **314,** 1137–1146.
36. Kanamaru, S., Leiman, P. G., Kostyuchenko, V. A., Chipman, P. R., Mesyanzhinov, V. V., Arisaka, F. & Rossmann, M. G. (2002) *Nature* **415,** 553–557.
37. Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R. & Reinemer, P. (1994) *Science* **265,** 383–386.
38. Yoder, M. D., Keen, N. T. & Jurnak, F. (1993) *Science* **260,** 1503–1507.
39. Huang, W., Matte, A., Li, Y., Kim, Y. S., Linhardt, R. J., Su, H. & Cygler, M. (1999) *J. Mol. Biol.* **294,** 1257–1269.
40. Lunin, V. V., Li, Y., Linhardt, R. J., Miyazono, H., Kyogashima, M., Kaneko, T., Bell, A. W. & Cygler, M. (2004) *J. Mol. Biol.* **337,** 367–386.
41. Thomassen, E., Gielen, G., Schutz, M., Schoehn, G., Abrahams, J. P., Miller, S. & van Raaij, M. J. (2003) *J. Mol. Biol.* **331,** 361–373.
42. Price, K. N., Tuinman, A., Baker, D. C., Chisena, C. & Cysyk, R. L. (1997) *Carbohydrate Res.* **303,** 303–311.
43. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950.

Title: The *Streptococcus equi* prophage-encoded protein SEQ2045 is a hyaluronan-specific hyaluronate lyase that is produced during equine infection

Running title: Hyaluronate lyase is produced during equine infection

Contents category: Biochemistry and Molecular Biology

Authors: Anna-Marie Lindsay,[1] Meng Zhang,[1] Zoe Mitchell,[2] Andrew S. Waller,[2] Iain C. Sutcliffe[1] and Gary W. Black[1]

Addresses: [1]Biomolecular and Biomedical Research Centre, School of Applied Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK; [2]Animal Health Trust, Lanwades Park, Kentford, Newmarket, Suffolk CB8 7UU, UK

Corresponding author: Gary W. Black, Biomolecular and Biomedical Research Centre, School of Applied Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK; Tel: +44 191 227 3519; Fax: +44 191 227 3550; Email: gary.black@northumbria.ac.uk.

List of abbreviations: Hyl, Hyaluronate lyase; HA, Hyaluronic acid.

Summary.

Streptococcus equi causes equine 'strangles'. Hyaluronate lyases, which degrade hyaluronan and chondroitins of connective tissue, are thought to facilitate streptococcal invasion of the host. However, prophage-encoded hyaluronate lyases are hyaluronan-specific and are thought to be primarily involved in the degradation of the hyaluronan capsule of streptococci during bacteriophage infection. To understand the role of prophage-encoded hyaluronate lyases further, we have biochemically characterised such a hyaluronate lyase, SEQ2045 from *S. equi*, and shown that it is produced during equine infection. Therefore, SEQ2045 may play a more direct role in disease pathogenesis than previously thought, although its primary role is most likely to be as a bacteriophage enzyme involved in the degradation of the hyaluronan capsule of *S. equi*.

Introduction.

Streptococcus equi is a Lancefield group C streptococcus. It is principally a pathogen of equines and causes the widespread and highly contagious disease 'strangles', which manifests as a nasopharyngeal infection that infiltrates the lymph nodes of the head and neck (Harrington *et al.*, 2002; Waller & Jolley, 2007). Regrettably, the development of an effective vaccine has been slow (Waller & Jolley, 2007). Hence a better understanding of *S. equi* pathogenesis is essential if improved vaccines are to be developed. Likely candidates for novel vaccine targets are the cell surface and extracellular virulence factors produced by *S. equi*.

The closely related human pathogen, *Streptococcus pyogenes* produces many virulence factors in common with *S. equi*, for example, superantigens (Proft *et al.*, 2003) and a hyaluronic acid (HA) capsule (Bisno *et al.*, 2003; Harrington *et al.*, 2002). Interestingly, several of these are encoded by genes carried by lysogenic bacteriophages that have integrated into the bacterial chromosome, i.e. prophages. One such prophage, SF370.1, from *S. pyogenes* SF370, encodes the pyrogenic mitogen SpeC and DNase virulence factors (Canchaya *et al.*, 2002). SF370.1 also encodes a hyaluronate lyase (Hyl), termed HylP1.

Typically, Hyls have been shown to be secreted and are capable of degrading HA and chondroitins, which are major components of mammalian connective tissue, thus possibly facilitating the invasion of the bacteria and their toxins (Hynes & Walton, 2000). However, HylP1 is thought not to be involved in the degradation of connective tissue, as it is unlikely to be secreted, since it lacks a signal peptide, and only degrades HA and does not process chondroitins (Smith *et al.*, 2005). Additionally, HylP1 is sequence-distinct to the Hyls that degrade chondroitins and has

been shown to contain a triple stranded β-helix; a structure only found in bacteriophage tail-fibre proteins (Smith *et al.*, 2005). It has therefore been postulated that the role of HylP1 is the penetration of the HA capsule of streptococci during phage infection (Hynes *et al.*, 1995), facilitating bacterial lysogenisation.

In an attempt to understand the role of prophage-encoded Hyls further, we have analysed SEQ2045, the closest HylP1 homologue in *S. equi*, as this organism also produces a HA capsule (Harrington *et al.*, 2002). We have shown that SEQ2045 is biochemically very similar to HylP1and that antibodies to SEQ2045 are produced by infected horses.

Methods.

2.1. Cloning

Genomic DNA of *S. equi* strain 4047 was isolated using a DNeasy extraction kit (Qiagen, UK). Synthetic oligonucleotides 5' CAT ATg TCA AAA gAA gTT gCA TC 3' and 5' ggA TCC TAT TTT TTT AgT ATg AgT TTT TTT AAC 3' were used to amplify *seq2045* from this genomic DNA using the KOD hot start polymerase kit (Merck Biosciences, UK). The amplified product was ligated into pCR-Blunt (Invitrogen, UK) then removed from pCR-Blunt on a *NdeI* - *XhoI* fragment and ligated into pET-28a (Merck Biosciences, UK) to produce pSEQ2045.

2.2. Protein expression

N-terminally hexahistidine-tagged SEQ2045 was produced in *Escherichia coli* BL21(DE3) (Merck Biosciences, UK) cultures carrying pSEQ2045 following the protocol described in (Charnock *et al.*, 2001), except cultures were grown at 30°C post-induction and concentration and exchange of the purified protein into 18.2 MΩ cm^{-1} H_2O was achieved using a 10 kDa cut-off concentrator unit (Viva Science, UK). The purity of SEQ2045 was judged by SDS/PAGE and Coomassie blue staining (Laemmli, 1970) and MS of a tryptic digest of the pure protein was used to confirm the identity of the protein according to the methods and criteria of (Zhang *et al.*, 2007).

2.3. Biochemical assays

Enzyme assays were performed as described by (Smith *et al.*, 2005) with reaction conditions of 2 mg ml^{-1} substrate; 8 mM $CaCl_2$; 50 mM CH_3COONH_4 (pH

7.5) unless otherwise stated. The range of HA concentrations used for determining k_{cat} and K_M values was 0.25 to 2 mg ml^{-1}. Goodness-of-fit statistical analysis of the linear trend lines of the resulting double-reciprocal plots produced R^2 values of $\geq$ 0.9042. The following buffers were used to determine the pH optimum; CH_3COONa (pH 5.0 and 5.5), NH_4Ac (pH 6.0, 6.5, 7.0 and 7.5) and glycine (pH 8.0). The effect of temperature on activity was determined by performing assays at 27, 37, 47, 57, 67 and 77°C. The temperature at which the rate of the reaction was reduced to 50% was determined by incubating assays at 27.5, 37.6, 47.4, 57.5, 67.6, 77.4 and 87.5°C for 20 min and then assaying at 37°C. High performance anion exchange chromatography was performed according to the method of (Lauder *et al.*, 2000).

2.4. Western blot analysis

Western blotting was performed as described by (Hamilton *et al.*, 2000) with 7 μg purified SEQ2045 per lane. The preimmune and strangles convalescent sera used were prepared at the Animal Health Trust. Control serum (donor herd serum) was purchased from Sigma Chemical Co.

Results.

SEQ2045 from S. equi is similar to HylP1 from S. pyogenes

A BLAST search of the *S. equi* strain 4047 genome (http://www.sanger.ac.uk/Projects/S_equi/) with ORF *hylP1* from *S. pyogenes* SF370 revealed a highly homologous ORF *seq2045* (E probability = 7.6 x 10^{-120}). The SEQ2045 and HylP1 proteins share 75% amino acid identity. BLASTP searches (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) with the sequences adjacent to *seq2045* in the genome of *S. equi* showed that *seq2045* resides within a prophage sequence that is very similar in its organisation to SF370.1, the prophage encoding HylP1 (Fig. 1). The *S. equi* prophage is notable as it also encodes proteins SEQ2036 (ORF #50) and SEQ2037 (ORF #49) (Fig. 1) that are 100% identical to the mitogens SeeH and SeeI from *S. equi* CF32 (Artiushin *et al.*, 2002).

Bioinformatic analyses predicted that SEQ2045 is a non-secretory protein (data not shown). We expressed full length recombinant *seq2045* in *E. coli* and its product, SEQ2045, was obtained in large quantities (~40 mg l^{-1}) as a 42 kDa *N*-terminally hexahistidine-tagged protein that was readily purified to homogeneity. Three non-redundant SEQ2045 peptides were identified by MS confirming the identity of the purified protein (Table 1). The temperature-activity optimum was shown to be 37°C (Fig. 2A) and the temperature at which the rate of the reaction was reduced by 50 % was shown to be 44°C (Fig. 2B). The pH–activity optimum for SEQ2045 against HA was shown to be ~7 (Fig. 2C). Kinetic analysis of SEQ2045 on HA substrate yielded a k_{cat} of 7.61 ± 0.63 s^{-1} and a K_M value of 0.29 ± 0.05 mg ml^{-1} (equivalent to 0.71 ± 0.13 mM, when the concentration of HA is expressed as the concentration of its disaccharide unit based on its molecular weight; Fig. 2D). The k_{cat}

values show a notable similarity to the k_{cat} determined for HylP1, 7.23 s^{-1}, and another Hyl, HylP, from bacteriophage H4489A from *S. pyogenes*, 4.9 s^{-1} (Baker *et al.*, 2002; Smith *et al.*, 2005). These turnover rates are in stark contrast to a sequence-distinct Hyl from *Streptococcus agalactiae* which has a k_{cat} of around 8600 s^{-1} on HA (Pritchard *et al.*, 1994). However, with regards to the binding capacity of SEQ2045 to HA, the calculated K_M value is ~5-fold lower than that of HylP1 (3.69 ± 0.06 mM) and ~8-fold higher than that of a sequence-distinct Hyl from *Streptococcus pneumoniae* (K_M value of 0.09 ± 0.03 mM) (Nukui *et al.*, 2003).

High performance anion exchange chromatography analysis of SEQ2045 digestion products of HA showed that end products were predominantly the Δ4,5 unsaturated hyalurono-hexasaccharide and the Δ4,5 unsaturated hyalurono-octasaccharide, with small amounts of Δ4,5 unsaturated hyalurono-tetrasaccharide, Δ4,5 unsaturated hyalurono-decasaccharide and longer oligosaccharides (Fig. 3). Early time points showed that a wide range of Δ4,5 unsaturated oligosaccharides are produced, indicating that SEQ2045 has an endo-acting mode of digestion. The wide size range of oligosaccharides released is the result of the random internal cuts of HA by SEQ2045.

Unlike sequence-distinct Hyls, SEQ2045 showed no catalytic activity against chondroitin 4-sulphate, chondroitin 6-sulphate and dermatin sulphate (data not shown).

SEQ2045 is produced during S. equi infection of horses

A *post facto* serological investigation of protein expression *in vivo* was undertaken using sera taken from a horse following natural *S. equi* infection (convalescent serum). Western blot analysis clearly showed that equine infection with

S. equi initiated an immunogenic response against SEQ2045 that was not observed with control sera (commercial horse sera from donor herds; data not shown). This indicated that specific antibodies are raised against SEQ2045 during the course of *S. equi* infection. Subsequently, results obtained with three pre- and post-infection matched pony sera samples (Fig. 4) confirmed that antibodies are made to SEQ2045 specifically in response to *S. equi* infection. These data are in excellent agreement with the findings of (Artiushin *et al.*, 2002), which showed that the prophage-encoded superantigens SeeH and SeeI (Fig. 1) are recognised by convalescent sera from horses that have recovered from *S. equi* infection.

Discussion.

SEQ2045 is a Hyl that is incapable of degrading HA further than a Δ4,5 unsaturated hyalurono-tetrasaccharide and has no activity against chondroitins. As such, it is unlikely that the principal role of the enzyme is the dissemination of *S. equi* through equine tissue and its primary role is most likely to be in phage penetration through the host capsule. However, it is clearly expressed during infection of horses, as demonstrated serologically, and it is notable that phage particles from lysates of *S. equi* have been demonstrated to possess *in vitro* Hyl activity (Timoney *et al.*, 1991). Thus phage induction *in vivo* may itself represent an additional mechanism leading to host connective tissue damage.

Interestingly, the only nonprophage-encoded Hyl of *S. equi* strain 4047, SEQ1479, is encoded by a pseudogene (our unpublished observations) and it has been noted previously that *in vitro* secreted Hyl activity was absent from culture supernatants of four out of ten *S. equi* clinical isolates (Sting *et al.*, 1990). Thus the extent to which *S. equi* is able to degrade host HA clearly varies from strain to strain. In some strains (e.g. strain 4047) the only source of Hyl activity is likely to be prophage-encoded, as prophage are apparently present in all *S. equi* strains (Alber *et al.*, 2004); our unpublished observations), whereas in other strains both prophage- and nonprophage-encoded enzymes may be present. It is noted that truncated Hyl sequences were also evident in a significant proportion of *Streptococcus suis* strains and that these strains lacked detectable Hyl activity (King *et al.*, 2004).

In light of the results presented here it may therefore be necessary to re-evaluate the role of HylP1 from *S. pyogenes* SF370. It is likely that HylP1 is expressed during human *S. pyogenes* infections as phage mobilisation has been

demonstrated both *in vitro* and in a mouse infection model (Broudy & Fischetti, 2003). HylP1 may therefore also act as an additional virulence factor to HylA, a sequence-distinct *S. pyogenes* secreted enzyme (SPY1032) which we have recently biochemically characterised and shown to be a Hyl with activity against HA and chondroitins (our unpublished observations). For both *S. pyogenes* and *S. equi*, phage induction *in vivo* may thus lead to the production and release of both phage-encoded virulence factors, e.g. DNAase (Broudy *et al.*, 2002) and superantigens, and phage structural proteins with the capacity to damage connective tissues (i.e. HylP1 and SEQ2045).

We have recently reported the importance of studying *S. equi* infections of the horse in the natural host (Hamilton *et al.*, 2006). Experimental infections of the horse with *S. equi* may therefore be a useful model for studying the role of Hyls in human *S. pyogenes* infections. Moreover, as the study of *in vivo* interactions between phages and pathogenic bacteria remains an important new frontier for phage biologists (Chibani-Chennoufi *et al.*, 2004), *S. equi* and the horse may provide an important system for studying the contribution of phages to bacterial virulence in the natural host.

References.

Alber, J., El-Sayed, A., Lammler, C., Hassan, A. A., Weiss, R. & Zschock, M. (2004). Multiplex polymerase chain reaction for identification and differentiation of Streptococcus equi subsp. zooepidemicus and Streptococcus equi subsp. equi. *J Vet Med B Infect Dis Vet Public Health* **51**, 455-458.

Artiushin, S. C., Timoney, J. F., Sheoran, A. S. & Muthupalani, S. K. (2002). Characterization and immunogenicity of pyrogenic mitogens SePE-H and SePE-I of Streptococcus equi. *Microb Pathog* **32**, 71-85.

Baker, J. R., Dong, S. & Pritchard, D. G. (2002). The hyaluronan lyase of Streptococcus pyogenes bacteriophage H4489A. *Biochem J* **365**, 317-322.

Bisno, A. L., Brito, M. O. & Collins, C. M. (2003). Molecular basis of group A streptococcal virulence. *Lancet Infect Dis* **3**, 191-200.

Broudy, T. B., Pancholi, V. & Fischetti, V. A. (2002). The in vitro interaction of Streptococcus pyogenes with human pharyngeal cells induces a phage-encoded extracellular DNase. *Infect Immun* **70**, 2805-2811.

Broudy, T. B. & Fischetti, V. A. (2003). In vivo lysogenic conversion of Tox(-) Streptococcus pyogenes to Tox(+) with Lysogenic Streptococci or free phage. *Infect Immun* **71**, 3782-3786.

Canchaya, C., Desiere, F., McShan, W. M., Ferretti, J. J., Parkhill, J. & Brussow, H. (2002). Genome analysis of an inducible prophage and prophage remnants integrated in the Streptococcus pyogenes strain SF370. *Virology* **302**, 245-258.

Charnock, S. J., Brown, I. E., Turkenburg, J. P., Black, G. W. & Davies, G. J. (2001). Characterization of a novel pectate lyase, Pel10A, from Pseudomonas cellulosa. *Acta Crystallogr D Biol Crystallogr* **57**, 1141-1143.

Chibani-Chennoufi, S., Bruttin, A., Dillmann, M. L. & Brussow, H. (2004). Phage-host interaction: an ecological perspective. *J Bacteriol* **186**, 3677-3686.

Hamilton, A., Harrington, D. & Sutcliffe, I. C. (2000). Characterization of acid phosphatase activities in the equine pathogen Streptococcus equi. *Syst Appl Microbiol* **23**, 325-329.

Hamilton, A., Robinson, C., Sutcliffe, I. C., Slater, J., Maskell, D. J., Davis-Poynter, N., Smith, K., Waller, A. & Harrington, D. J. (2006). Mutation of the maturase lipoprotein attenuates the virulence of Streptococcus equi to a greater extent than does loss of general lipoprotein lipidation. *Infect Immun* **74**, 6907-6919.

Harrington, D. J., Sutcliffe, I. C. & Chanter, N. (2002). The molecular basis of Streptococcus equi infection and disease. *Microbes Infect* **4**, 501-510.

Hynes, W. L., Hancock, L. & Ferretti, J. J. (1995). Analysis of a second bacteriophage hyaluronidase gene from Streptococcus pyogenes: evidence for a third

hyaluronidase involved in extracellular enzymatic activity. *Infect Immun* **63**, 3015-3020.

Hynes, W. L. & Walton, S. L. (2000). Hyaluronidases of Gram-positive bacteria. *FEMS microbiology letters* **183**, 201-207.

King, S. J., Allen, A. G., Maskell, D. J., Dowson, C. G. & Whatmore, A. M. (2004). Distribution, genetic diversity, and variable expression of the gene encoding hyaluronate lyase within the Streptococcus suis population. *J Bacteriol* **186**, 4740-4747.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lauder, R. M., Huckerby, T. N. & Nieduszynski, I. A. (2000). A fingerprinting method for chondroitin/dermatan sulfate and hyaluronan oligosaccharides. *Glycobiology* **10**, 393-401.

Nukui, M., Taylor, K. B., McPherson, D. T., Shigenaga, M. K. & Jedrzejas, M. J. (2003). The function of hydrophobic residues in the catalytic cleft of Streptococcus pneumoniae hyaluronate lyase. Kinetic characterization of mutant enzyme forms. *J Biol Chem* **278**, 3079-3088.

Pritchard, D. G., Lin, B., Willingham, T. R. & Baker, J. R. (1994). Characterization of the group B streptococcal hyaluronate lyase. *Arch Biochem Biophys* **315**, 431-437.

Proft, T., Webb, P. D., Handley, V. & Fraser, J. D. (2003). Two novel superantigens found in both group A and group C Streptococcus. *Infect Immun* **71**, 1361-1369.

Smith, N. L., Taylor, E. J., Lindsay, A. M., Charnock, S. J., Turkenburg, J. P., Dodson, E. J., Davies, G. J. & Black, G. W. (2005). Structure of a group A streptococcal phage-encoded virulence factor reveals a catalytically active triple-stranded beta-helix. *Proc Natl Acad Sci U S A* **102**, 17652-17657.

Sting, R., Schaufuss, P. & Blobel, H. (1990). Isolation and characterization of hyaluronidases from Streptococcus dysgalactiae, S. zooepidemicus and S. equi. *Zentralbl Bakteriol* **272**, 276-282.

Timoney, J., Pesante, L. & Ernst, C. (1991). Hyaluronidase associated with temperate bacteriophage of Streptococcus equi. In *Genetics and Molecular Biology of Streptococci, Lactococci and Enterococci* Edited by G. Dunny, P. Cleary & L. McKay. Washington DC: American Society for Microbiology.

Waller, A. S. & Jolley, K. A. (2007). Getting a grip on strangles: recent progress towards improved diagnostics and vaccines. *Vet J* **173**, 492-501.

Zhang, M., McDonald, F. M., Sturrock, S. S., Charnock, S. J., Humphery-Smith, I. & Black, G. W. (2007). Group A streptococcus cell-associated pathogenic proteins as revealed by growth in hyaluronic acid-enriched media. *Proteomics* 7, 1379-1390.

Figure legends.

Fig. 1. The gene maps of the *S. equi* 4047 prophage that carries ORF *seq2045* and the *S. pyogenes* prophage SF370.1. ORFs are not depicted where they are less than 60 codons in length and/or when blastp returned a result of "No significant similarity found". The *S. equi* prophage ORFs (SEQ2089-SEQ2035) are numbered from 1 to 51. The *S. pyogenes* ORFs were numbered 1 to 48 (Canchaya *et al.*, 2002). Genes likely to belong to the same module were marked with the same pattern. The mitogens SEQ2036 (SeeH) and SEQ2037 (SeeI) (Artiushin *et al.*, 2002) correspond to ORF #50 and #49 respectively.

Fig. 2. Graphs used in the determination of the biochemical and biophysical properties of SEQ2045. A, plot of temperature against rate of reaction for temperature-activity optimum determination; B, plot of pre-incubation temperature against rate of reaction at 37 ^{o}C for determining the temperature at which the rate of the reaction was reduced to 50 % (i.e. thermal midpoint); C, plot of pH against rate of reaction for the pH-activity optimum determination; and D, a double-reciprocal plot, the insert is the corresponding non-reciprocal plot.

Fig. 3. High performance anion exchange chromatograms of the products of a digestion of HA with SEQ2045. A, 0 min digestion; B, 10 min digestion; C, 16 h digestion. The degree of polymerization of product is shown. mAU, milli absorbance units.

Fig. 4. Western blot showing the immunogenic response of SEQ2045 to sera from ponies prior to and subsequent to natural infection with *S. equi*. 1, pre-infection sera for Pony #1; 2, post-infection sera for Pony #1; 3, Precision Plus Protein standards (Bio-Rad, UK); 4, pre-infection sera for Pony #2; 5, post-infection sera for Pony #2; 6, pre-infection sera for Pony #3; 7, post-infection sera for Pony #3.

Tables and figures.

Table 1. SEQ2045 peptides identified by MS.

Peptide sequence	Charge state	X_{corr} *	ΔC_n †	Location‡
ETSQIDGNLK	2	3.15	0.47	333 - 342
ALGTLK	1	1.60	0.26	255 - 260
KQKGGK	1	1.77	0.21	286 - 291

* cross correlation score
† difference in normalised cross correlation score
‡ Percent amino acid coverage of entire protein of the three peptide sequences is 5.91

Fig. 1

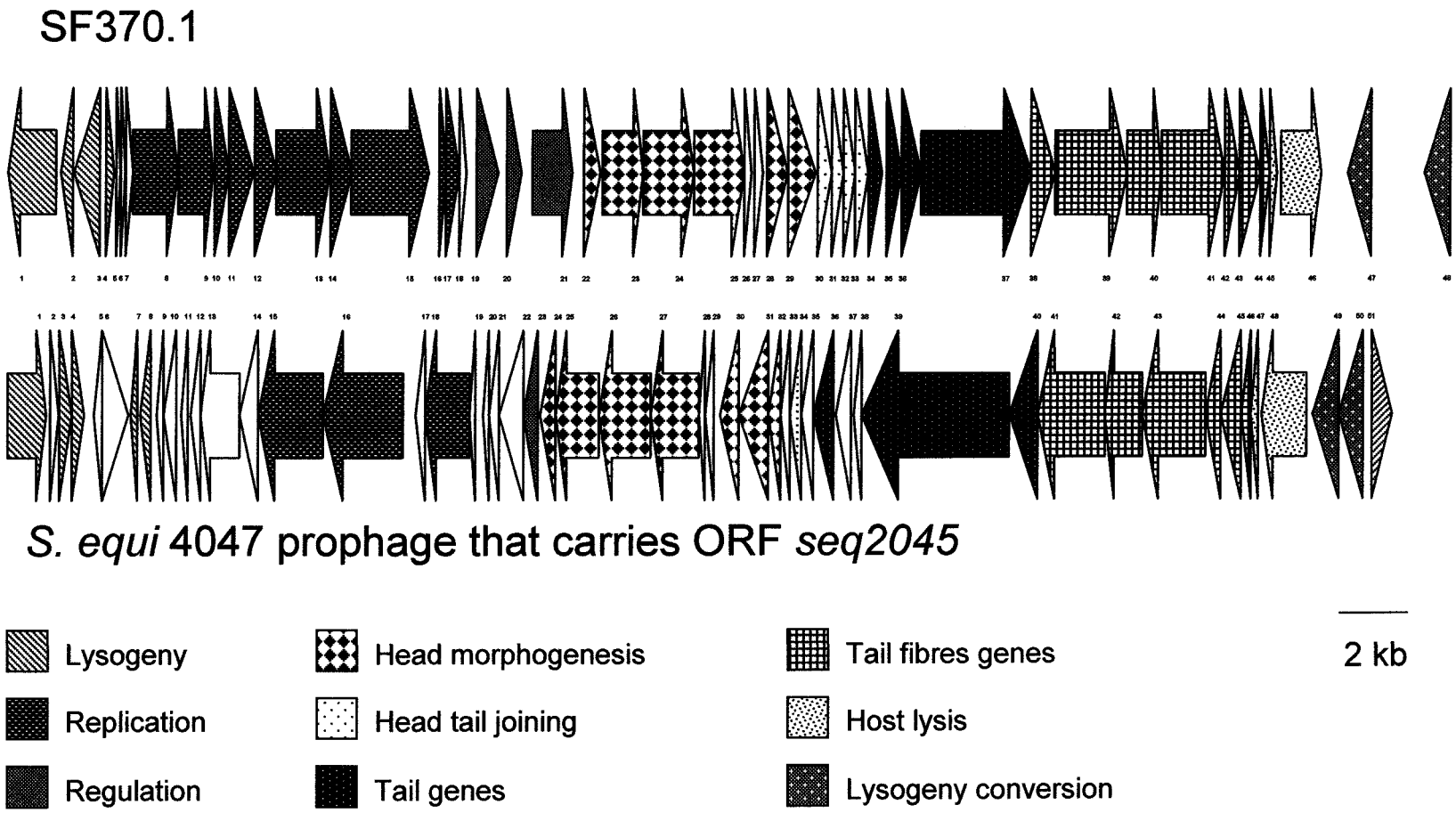

SF370.1
S. equi 4047 prophage that carries ORF seq2045
Lysogeny
Replication
Regulation
Head morphogenesis
Head tail joining
Tail genes
Tail fibres genes
Host lysis
Lysogeny conversion
2 kb

Fig. 2

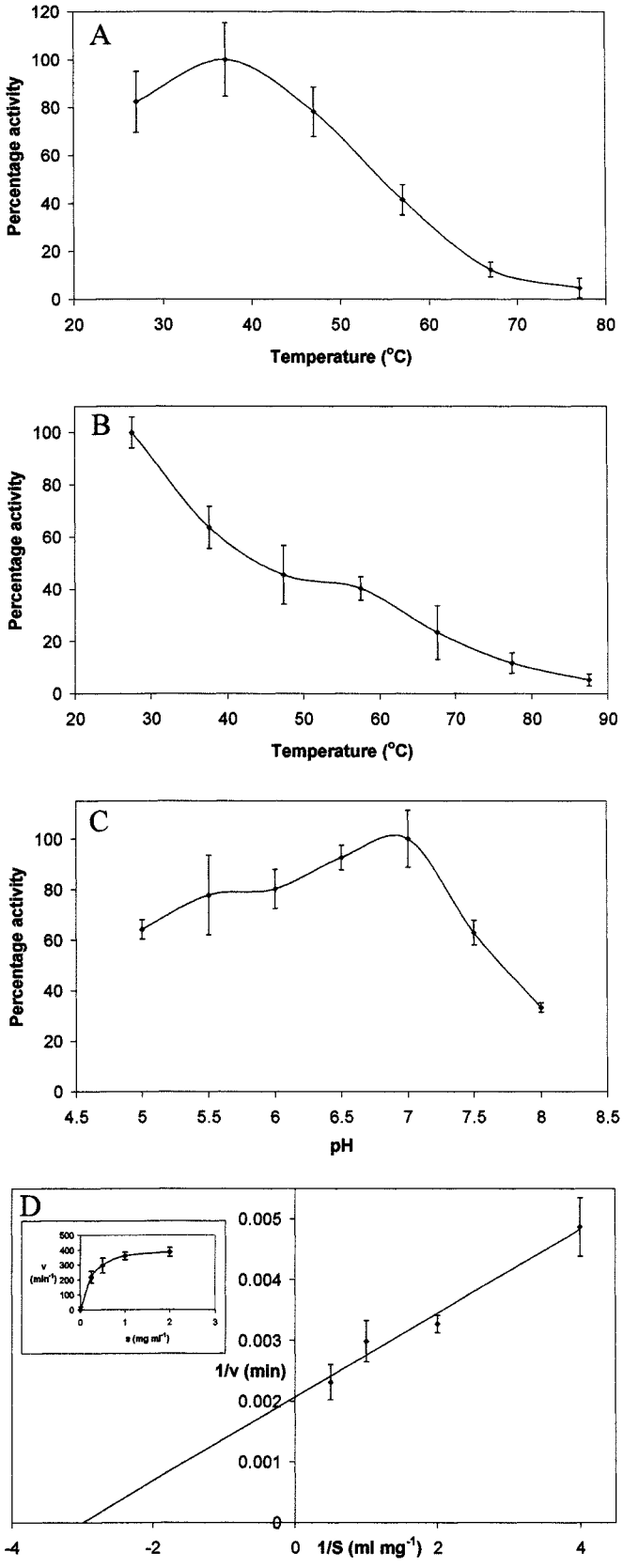

A
Percentage activity
Temperature (°C)
B
Percentage activity
Temperature (°C)
C
Percentage activity
pH
D
1/v (min)
1/S (ml mg-1)

Fig. 3

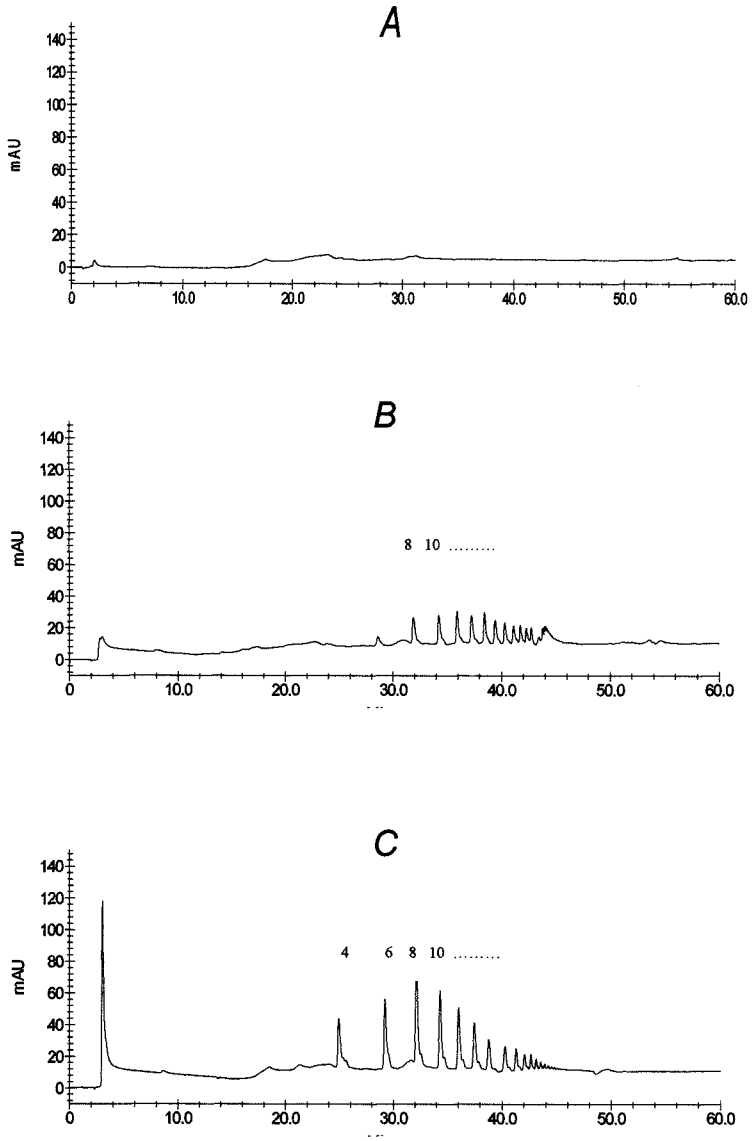
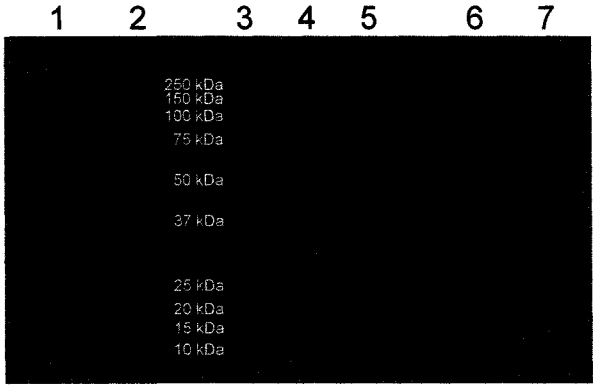

A
mAU
140
120
100
80
60
40
20
0
0
10.0
20.0
30.0
40.0
50.0
60.0
B
mAU
8 10 .....
C
mAU
4 6 8 10 .....

Fig. 4


1
2
3
4
5
6
7
250 kDa
150 kDa
100 kDa
75 kDa
50 kDa
37 kDa
25 kDa
20 kDa
15 kDa
10 kDa